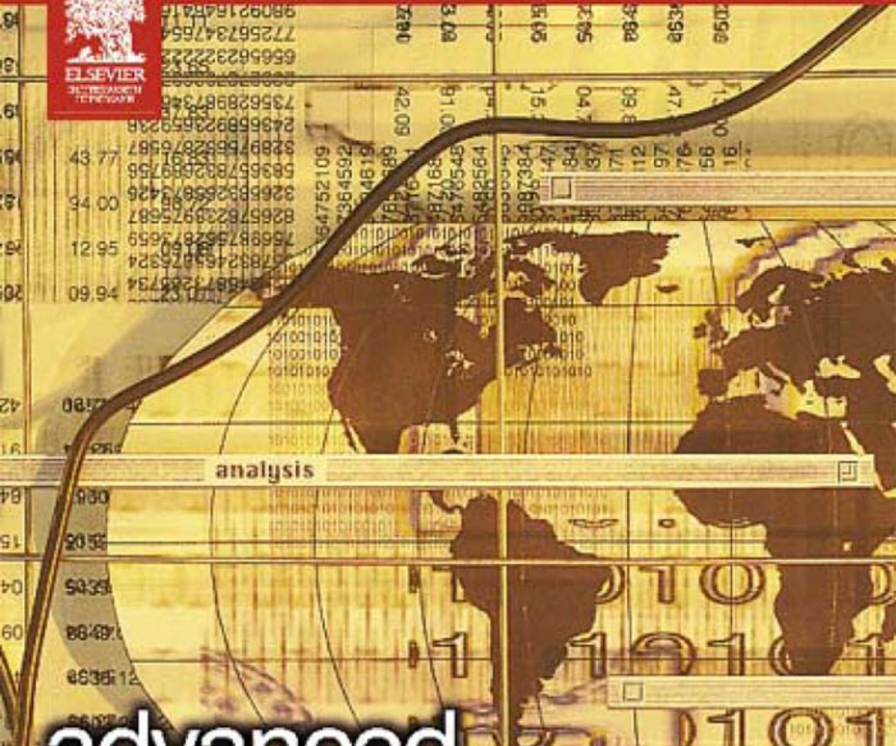
ELSEVIER FINANCE



advanced fixed income analysis

Moorad Choudhry

Foreword by Professor Christine Oughton
Birkbeck, University of London

Advanced Fixed Income Analysis

This page intentionally left blank

Advanced Fixed Income Analysis

Moorad Choudhry

ELSEVIER
BUTTERWORTH
HEINEMANN

AMSTERDAM BOSTON HEIDELBERG LONDON NEW YORK OXFORD
PARIS SAN DIEGO SAN FRANCISCO SINGAPORE SYDNEY TOKYO

Elsevier Butterworth-Heinemann
Linacre House, Jordan Hill, Oxford OX2 8DP
200 Wheeler Road, Burlington, MA 01803

First published 2004

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication Data
A catalogue record for this book is available from the Library of Congress

ISBN 0 7506 6263 8
0 7506 6279 4 (Fixed Income Markets Library)

For information on all Elsevier Butterworth-Heinemann finance publications visit our website at http://books.elsevier.com/finance

Typeset by Integra Software Services Pvt. Ltd, Pondicherry, India
www.integra-india.com
Printed and bound in Great Britain

For my grandfather, Mr Abdul Hakim (c. 1898–1983), Advocate.
Citizen of Noakhali, Bangladesh

Contents

Foreword *ix*

About the Author *xi*

Preface *xiii*

1 Approaches to Trading and Hedging 1

2 Relative Value Trading using Government Bonds 19

3 The Dynamics of Asset Prices 35

4 Interest-rate Models I 55

5 Interest-rate Models II 79

6 Estimating and Fitting the Term Structure I 97

7 Estimating and Interpreting the Term Structure II: a Practical Implementation of the Cubic Spline Method 117

8 Advanced Analytics for Index-linked Bonds 141

9 Analysing the Long Bond Yield 153

10 The Default Risk of Corporate Bonds 163

11 Brady Bonds 169

Index *175*

Foreword

The fixed income markets have always been centres of innovation and creativity. This much is apparent from even a cursory glance at developments in recent and not-so-recent history. However, it is only in the last twenty years or so that such innovation has really been required, as markets changed significantly and capital started to move freely. The bond market has been the vital conduit through which capital has been raised; continuing product development in the markets has made a significant, and irreplaceable, contribution to global economic progress. The range of products available is vast and growing, as the needs of both providers and users of capital continually alters in response to changing conditions. This economic dynamic means that market participants observe a state of constant learning, as they must if they are to remain effective in their work. Consider, for instance, the new instruments and techniques that we have had to become familiar with in just the last few years: new instruments for hedging credit risk, new techniques for raising capital through synthetic securitisation of the most esoteric 'reference' assets, and new models for fitting the term structure of interest rates – there is much for market participants to keep in touch with. Inevitably practitioners are required to become specialists, as each segment of the debt markets demands increasingly complex approaches in addressing its problems and requirements.

Of course, users of capital are not limited to existing products for raising finance or hedging market risk exposure. They can ask an investment bank to design an instrument to meet their individual requirements, and target it at specific groups of customers. For example, it is arguable whether the growth of the 'credit-card banks' in the United States (such as MBNA) could have occurred so rapidly without the securitisation mechanism that enabled them to raise lower-cost funding. Witness also the introduction of exotic structured credit products, such as the synthetic collateralised debt obligation (CDO), which uses credit derivatives in its construction and followed rapidly on the development of more conventional CDO structures. The so-called 'CSO' was designed to meet regulatory capital and credit risk management requirements, as opposed to funding requirements. The increasing depth and complexity of the markets requires participants to be completely up-to-date on the latest analytical and valuation techniques if they are not to risk being left behind. It is clear that we operate in an environment in which there exists a long-term interest in the application of ever more sophisticated valuation and analytical techniques. The level of mathematical sophistication in use in financial markets today is phenomenal, not to mention very specialised.

That is why this book, from one of the leading researchers and writers on fixed income today, is such a welcome publication. I should of course say 'books', as we have a series here that forms part of a handsome Library. The antecedents of the author promise that these books will make a high-quality contribution to the field. But it is the books' clarity of approach and focus that I am most excited about. The books are welcome because they are part of the continuing need to remain, as Alan Greenspan would have said, ahead of the curve. They contain insights into practical techniques and applications used in the fixed income markets today, with a hint at what one might expect in the future. They also indicate the scope and significance of these techniques in the world of finance. Readers will notice that the text is fairly technical at many points. This reflects the level of mathematical sophistication one encounters in the markets.

If the author will indulge me, I would like to highlight those parts of the books I was particularly interested in.

The treatment of yield curve analysis in *Advanced Fixed Income Analysis* is first rate. For instance, I liked the comprehensive description of the 'variable roughness penalty' approach to cubic spline estimation of the term structure (Chapter 6). The author rightly points out that most market practitioners can have their analytical needs met by the simpler techniques of yield curve fitting, and only exotic option traders, who wish to model the volatility surface, really need to resort to multi-factor term structure models. That is why the practical demonstration of the cubic spline technique is so welcome in this book. Portfolio managers using this technique will get a good understanding of recent movements in the yield curve as well as good interpretive information for the future. Elsewhere we have a comprehensive treatment of the main single-factor and multi-factor yield curve models in use, with useful comment on the efficacies of using both. The practical implications of using the different interest-rate models are well handled and Chapters 4 and 5 will be of value to practitioners. There is also accessible coverage of the Heath–Jarrow–Morton interest-rate model, described and explained here in its single-factor and multi-factor forms. The author cleverly draws out the link between academic research and market applications by showing how financial institutions are able to continue meeting their clients' ever more complex requirements by incorporating insights from research into their product development.

I am very enthusiastic about the book *Corporate Bonds and Structured Financial Products*. The author captures all the key capital raising instruments. I was fascinated to learn about the synthetic asset-backed CP structure or 'conduit'. Distinct from conventional AB-CP programmes, I was very interested to read about this. Of course, one might (in hindsight!) easily have predicted its development, mirroring as it did the practice seen in the bonds and note market when credit derivatives were allied with traditional securitisation techniques to produce the synthetic CDO. The author presents a new look at established and new products, and both venerable and brand-new techniques. As such the book should be practical interest to fund managers and traders, as well as corporate treasurers.

It is a privilege to be asked to write this foreword. By drawing on both his practical experience of financial markets and research for his PhD at Birkbeck, University of London, Moorad Choudhry successfully combines insights from theory and practice to make a genuinely worthwhile contribution to the financial economics literature. I do hope that this exciting and interesting new Library spurs readers on to their own research and investigation; if they follow the application and dedication evident in this work, they will not be going far wrong.

Professor Christine Oughton
School of Management and Organizational Psychology
Birkbeck, University of London
March 2004

About the Author

Moorad Choudhry is Head of Treasury at KBC Financial Products in London. He previously worked in structured finance services at JPMorgan Chase Bank, and as a government bond trader at Hambros Bank Limited and ABN Amro Hoare Govett Sterling Bonds Limited.

Dr Choudhry is a Visiting Professor at the Department of Economics, London Metropolitan University, a Senior Fellow at the Centre for Mathematical Trading and Finance, CASS Business School, and a Fellow of the Securities Institute.

Map of Surrey, showing area of Outstanding Natural Beauty.
Drawn from information supplied by The Surrey
Hills Partnership. Reproduced with permission.

Moorad Choudhry lives in Surrey, England.

Preface

This book is a review of the bond markets but with specific emphasis on the fundamental questions of the yield curve, the processes that determine the evolution of the interest-rate term structure, and how these processes can be modelled. We consider some advanced topics in bond market theory, so readers should be familiar with the basic concepts including interest-rate risk, convexity, the pricing process for bonds with embedded options and the concept of option-adjusted spread, the fundamentals of index-linked bonds, and the application of the Black–Scholes model in options pricing. This book reviews:

- the concept of relative value trading using government bonds;
- a number of interest-rate models and the assumptions underlying these models.

We also look at the procedure involved in estimating and fitting the yield curve. This topic is one of the most heavily researched subjects in financial market economics, and indeed research is ongoing. There are a number of ways to estimate and fit the yield curve, and there is no one right or wrong method.

It is important, both when discussing the subject matter or writing about it, to remember to place the relevant ideas in context, otherwise there is the danger of becoming too theoretical. The aim is to confine the discussion within the boundary of user application; there is a great deal of published material that is, quite simply, rather too theoretical, not to mention highly technical. We must always try to keep in touch with the markets themselves. The chapters are written from the point of view of both the market practitioner and the research student.

In the following chapters we look at bond relative value trading as well as interest-rate modelling. We summarise some interest-rate models, in a practical way that excludes most of the mathematics. The aim is firmly to discuss application of the models, and not to derive them or prove the maths. In this way readers should be able to assess the different methodologies for themselves and decide the efficacies of each for their own purposes. As always, selected recommended texts, plus chapter references are listed at the back, and would be ideal as a starting point for further research. This serves to highlight that this book is very much a summary of the latest developments, rather than a fully comprehensive review of the subject. The topics would be suitable for a separate book in their own right, and such a book would make an ideal companion to this book. However there is sufficient detail and exposition here to leave the reader with, hopefully, a good understanding of the subject.

To begin with we look at relative value trading, and some aspects of this for the bond trader. We then move on to the second part of this book, and introduce the dynamics of asset pricing, which is fundamental to an understanding of yield curve analysis. We review the main one-factor models that were initially developed to model the term structure. This includes the Vasicek, Cox–Ingersoll–Ross and Hull–White models. In most cases the model result is given and explained, rather than the full derivation. The objective here is to keep the content accessible, and pertinent to practitioners and most postgraduate students. A subsequent book is planned that will delve deeper into the models themselves, and the latest developments in research. Later we look at more advanced multi-factor models, led by the Heath–Jarrow–Morton model. Finally, we review some techniques used to estimate and fit the zero-coupon curve using the prices of bonds observed in the market, with an illustration from the United Kingdom gilt market.

The last part of the book considers some advanced analytical techniques for index-linked bonds. Chapter 9 is a look at some of the peculiar properties of very long-dated bond yields, including the convexity bias inherent in such yields, and their relative volatility. In Chapter 10 we review some concepts that apply to the analysis of the credit default risk of corporate bonds, and how this might be priced.

The dynamics of the yield curve

In Chapter 2 of the companion volume to this book in the Fixed Income Markets Library, *Corporate Bonds and Structured Financial Products*, we introduced the concept of the yield curve, and reviewed some preliminary issues concerning both the shape of the curve and to what extent the curve could be used to infer the shape and level of the yield curve in the future. We do not know what interest rates will be in the future, but given a set of zero-coupon (spot) rates today we can estimate the future level of forward rates (given today's spot rates) using a yield curve model. In many cases, however, we do not have a zero-coupon curve to begin with, so it then becomes necessary to derive the spot yield curve from the yields of coupon bonds, which one can observe readily in the market. If a market only trades short-dated debt instruments, then it will be possible to construct a short-dated spot curve.

It is important for a zero-coupon yield curve to be constructed as accurately as possible. This is because the curve is used in the valuation of a wide range of instruments, not only conventional cash market coupon bonds, which we can value using the appropriate spot rate for each cash flow, but other interest-rate products such as swaps.

If using a spot rate curve for valuation purposes, banks use what are known as *arbitrage-free* yield curve models, where the derived curve has been matched to the current spot yield curve. The concept of arbitrage-free, also known as no-arbitrage pricing or 'the law of one price' is that if one is valuing the same product or cash flow in two different ways, the same result will be obtained from either method. So, if one was valuing a two-year bond that was put-able by the holder at par in one year's time, it could be analysed as a one-year bond that entitled the holder to reinvest it for another year. The rule of no-arbitrage pricing states that an identical price will be obtained whichever way one chooses to analyse the bond. When matching derived yield curves therefore, correctly matched curves will generate the same price when valuing a bond, whether a derived spot curve is used or the current term structure of spot rates.

From our understanding of derivatives, we know that option pricing models such as Black–Scholes assume that asset price returns follow a lognormal distribution. The dynamics of interest rates and the term structure is the subject of some debate, and the main difference between the main interest-rate models is in the way that they choose to capture the change in rates over a time period. However, although volatility of the yield curve is indeed the main area of difference, certain models are easier to implement than others, and this is a key factor a bank considers when deciding which model to use. The process of *calibrating* the model, that is setting it up to estimate the spot and forward term structure using current interest rates that are input to the model, is almost as important as deriving the model itself. So the availability of data for a range of products, including cash money markets, cash bonds, futures and swaps, is vital to the successful implementation of the model.

As one might expect the yields on bonds are correlated, in most cases very closely positively correlated. This enables us to analyse interest-rate risk in a portfolio for example, but also to model the term structure in a systematic way. Much of the traditional approach to bond portfolio management assumed a parallel shift in the yield curve, so that if the 5-year bond yield moved upwards by 10 basis points, then the 30-year bond yield would also move up by 10 basis points. This underpins traditional duration and modified duration analysis, and the concept of immunisation. To analyse bonds in this way, we assume therefore that bond yield volatilities are identical and correlations are perfectly positive. Although both types of analysis are still common, it is clear that bond yields do not move in this fashion, and so we must enhance our approach in order to perform more accurate analysis.

Factors influencing the yield curve

From the discussion in Chapter 2 of the companion volume to this book in the Fixed Income Markets Library, *Corporate Bonds and Structured Financial Products* we are aware that there are a range of factors that impact on the shape and level of the yield curve. A combination of economic and non-economic factors are involved. A key factor is investor expectations, with respect to the level of inflation, and the level of real interest rates in the future. In the real world the market does not assume that either of these two factors is constant, however given that there is a high level of uncertainty over anything longer than the short-term, generally there is an assumption about both inflation and interest rates to move towards some form of equilibrium in the long-term.

It is possible to infer market expectations about the level of real interest rates going forward by observing yields in government index-linked bonds, which trade in a number of countries including the US and UK. The market's view on the future level of interest rates may also be inferred from the shape and level of the current yield curve. We know that the slope of the yield curve also has an information content. There is more than one way to interpret any given slope, however, and this debate is still open.

The fact that there are a number of factors that influence changes in interest rates and the shape of the yield curve means that it is not straightforward to model the curve itself. In Chapter 6 we consider some of the traditional and more recent approaches that have been developed.

Approaches to modelling

The area of interest-rate dynamics and yield curve modelling is one of the most heavily researched in financial economics. There are a number of models available in the market today, and generally it is possible to categorise them as following certain methodologies. By categorising them in this way, participants in the market can assess them for their suitability, as well as draw their own conclusions about how realistic they might be. Let us consider the main categories.

One-factor, two-factor and multi-factor models

The key assumption that is made by an interest-rate model is whether it is one-factor, that is the dynamics of the yield change process are based on one factor, or multi-factor. From observation we know that in reality there are a number of factors that influence the price change process, and that if we are using a model to value an option product, the valuation of that product is dependent on more than one underlying factor. For example, the payoff on a bond option is related to the underlying bond's cash flows as well as to the reinvestment rate that would be applied to each cash flow, in addition to certain other factors. Valuing an option therefore is a multi-factor issue. In many cases, however, there is a close degree of correlation between the different factors involved. If we are modelling the term structure, we can calculate the correlation between the different maturity spot rates by using a covariance matrix of changes for each of the spot rates, and thus obtain a common factor that impacts all spot rates in the same direction. This factor can then be used to model the entire term structure in a one-factor model, and although two-factor and multi-factor models have been developed, the one-factor model is still commonly used. In principle it is relatively straightforward to move from a one-factor to a multi-factor model, but implementing and calibrating a multi-factor model is a more involved process. This is because the model requires the estimation of more volatility and correlation parameters, which slows down the process.

Readers will encounter the term *Gaussian* in reference to certain interest-rate models. Put simply, a Gaussian process describes one that follows a normal distribution under a probability density function. The distribution of rates in this way for Gaussian models implies that interest rates can attain negative values under positive probability, which makes the models undesirable for some market practitioners. Nevertheless, such models are popular because they are relatively straightforward to implement and because the probability of the model generating negative rates is low and occurs only under certain extreme circumstances.

The short-term rate and the yield curve

The application of risk-neutral valuation requires that we know the sequence of short-term rates for each scenario, which is provided in some interest-rate models. For this reason, many yield curve models are essentially models of the stochastic evolution of the short-term rate. They assume that changes in the short-term interest rate is a *Markov* process. (It is outside the scope of this book to review the mathematics of such processes, but references are provided in subsequent chapters.) This describes an evolution of short-term rates in which the evolution of the rate is a function only of its current level, and not the path by which it arrived there. The practical significance of this is that the valuation of interest-rate products can be reduced to the solution of a single partial differential equation.

Short-rate models are composed of two components. The first attempts to capture the average rate of change, also called the *drift*, of the short-term rate at each instant, while the second component measures this drift as a function of the volatility of the short-term rate. This is given by:

$$\mathrm{d}r(t) = \mu(r,\ t)\mathrm{d}t + \sigma(r,\ t)\mathrm{d}W(t)$$

where $\mathrm{d}r(t)$ is the instantaneous change in the short-term rate, and $W(t)$ is the stochastic process that describes the evolution in interest rates, known as a Brownian or Weiner process.

The term $\mu(r,\ t)$ is the value of the drift multiplied by the size of the time period. The term $\sigma(r,\ t)\mathrm{d}W(t)$ is the volatility of the short-term rate multiplied by a random increment that is normally distributed. In most models the drift rate term is determined through a numerical technique that matches the initial spot rate yield curve, while in some models an analytical solution is available. Generally models assume an arbitrage-free relationship between the initial forward rate curve, the volatility $\sigma(r, t)\mathrm{d}W(t)$, the market price of interest-rate risk and the drift term $\mu(r, t)$. In models such as those presented by Vasicek (1977) and Cox–Ingersoll–Ross (1985), the initial spot rate yield curve is given by an analytical formula in terms of the model parameters, and they are known as *equilibrium* models, because they describe yield curves as being derived from an assumption of economic equilibrium, based on a given market interest rate. So the Vasicek and CIR models are models of the short-term rate, and both incorporate the same form for the drift term, which is a tendency for the short-term rate to rise when it is below the long-term mean interest rate, and to fall when it is above the long-term mean. This is known as *mean reversion*. Therefore we can describe the short-term drift rate in the form:

$$\mu = \kappa(\theta - r)$$

where r is the short-term rate as before and κ and θ are the mean reversion and long-term rate constants. In the Vasicek model, the rate dependence of the volatility is constant, in the CIR model it is proportional to the square-root of the short rate. In both models, because the dynamics of the short-rate cover all possible moves, it is possible to derive negative interest rates, although under most conditions of initial spot rate and volatility levels, this is quite rare. Essentially the Vasicek and CIR models express the complete forward rate curve as a function of the current short-term rate, which is why later models are sometimes preferred.

Other models that are similar in concept are the Black–Derman–Toy (1990) and Black–Karinski (1992) models, however these have different terms for the drift rate and require numerical fitting to the initial interest rate and volatility term structures. The drift rate term is not known analytically in these models. In the BDT model the short-term rate volatility is related to the strength of the mean reversion in a way that reduces the volatility over time.

Arbitrage-free and equilibrium modelling

In an arbitrage-free model, the initial term structure described by spot rates today is an input to the model. In fact such models could be described not as models per se, but essentially a description of an arbitrary process that governs changes in the yield curve, and projects a forward curve that results from the mean and volatility of the current short-term rate. An equilibrium term structure model is rather more a true model of the term structure process; in an equilibrium model the current term structure is an output from the model. An equilibrium model employs a statistical approach, assuming that market prices are observed with some statistical error, so that the term structure must be estimated, rather than taken as given.

Risk-neutral probabilities

When valuing an option written on say, an equity the price of the underlying asset is the current price of the equity. When pricing an interest-rate option the underlying is obtained via a random process that describes the instantaneous risk-free zero-coupon rate, which is generally termed the short rate.

In the following chapters we explore the different models that may be used and their application.

Mathematics primer

The level of mathematics required for a full understanding of even intermediate concepts in finance is frighteningly high. To attempt to summarise even the basic concepts in just a few pages would be a futile task and might give the impression that the mathematics was being trivialised. Our intention is quite the opposite. As this is a financial markets book, and not a mathematics textbook, a certain level of knowledge has been assumed, and a formal or rigorous approach has not been adopted. Hence readers will find few derivations, and fewer proofs. What we provide here is a very brief introduction to some of the concepts; the aim of this is simply to provide a starting point for individual research. We assist this start by listing recommended texts in the bibliography.

Random variables and probability distributions

In financial mathematics random variables are used to describe the movement of asset prices, and assuming certain properties about the process followed by asset prices allows us to state what the expected outcome of events are. A random variable may be any value from a specified *sample space*. The specification of the *probability distribution* that applies to the sample space will define the frequency of particular values taken by the random variable. The cumulative distribution function of a random variable X is defined using the distribution function $f(\ )$ such that $Pr\{X \leq x\} = f(\ )$. A discrete random variable is one that can assume a finite or *countable* set of values, usually assumed to be the set of positive integers. We define a discrete random variable X with its own probability function $p(i)$ such that $p(i) = Pr\{X = i\}$. In this case the probability distribution is

$$f(i) = Pr\{X \leq i\} = \sum_{n=0}^{i} p(n)$$

with $0 \leq p(i) \leq 1$ for all i. The sum of the probabilities is 1.

Discrete probability distributions include the Binomial distribution and the Poisson distribution.

Continuous random variables

The next step is to move to a continuous framework. A continuous random variable X may assume any real value and its probability density function $f(x)$ is defined as

$$f(x) = \lim_{\mathrm{d}x \to 0} \frac{Pr\{x \leq X \leq x + \mathrm{d}x\}}{\mathrm{d}x} = \frac{\mathrm{d}F(x)}{\mathrm{d}x}.$$

The probability distribution function is given as $F(x) = Pr\{X \leq x\} = \int_{s=-\infty}^{x} f(s)\mathrm{d}s$.

Continuous distributions are commonly encountered in finance theory. The *normal* or Gaussian distribution is perhaps the most important. It is described by its mean μ and standard deviation σ, sometimes called the location and spread respectively. The probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

Where a random variable X is assumed to follow a normal distribution it will be described in the form $X \sim N(\mu, \sigma^2)$ where $\sim$ means 'is distributed according to'. The standard normal distribution is written as $N(0, 1)$ with $\mu = 0$ and $\sigma = 1$. The cumulative distribution function for the standard normal distribution is given by

$$\Phi(x) = N(x) = \int_{z=-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \mathrm{d}z.$$

The key assumption in the derivation of the Black–Scholes option pricing model is that the asset price follows a *lognormal* distribution, so that if we assume the asset price is P we write

$$log\left(\frac{P_t}{P_0}\right) \sim N((r - \tfrac{1}{2}\sigma^2)t,\ \sigma^2 t).$$

Expected values

A probability distribution function describes the distribution of a random variable X. The expected value of X in a discrete environment is given by

$$E[X] = \overline{X} = \sum_{i=0}^{\infty} ip(i)$$

and the equivalent for a continuous random variable is

$$E[X] = \overline{X} = \int_{s=-\infty}^{\infty} sf(s)\mathrm{d}s.$$

The dispersion around the mean is given by the *variance* which is

$$Var[X] = E(X - \overline{X})^2 = \sum_{i=0}^{\infty} (i - \overline{X})^2\ p(i)$$

or

$$Var[X] = E(X - \overline{X})^2 = \int_{s=-\infty}^{\infty} f(s)\mathrm{d}s$$

in a continuous distribution. A squared measure has little application so commonly the square root of the variance, the standard deviation is used.

Regression analysis

A linear relationship between two variables, one of which is dependent, can be estimated using the least squares method. The relationship is

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

where X is the independent variable and ε is an error term capturing those explanatory factors not covered by the model. ε is described as $\varepsilon_i \sim N(0, \sigma^2)$. β is the slope of the linear regression line that describes the relationship, while α is the intercept of the y-axis. The sum of the squares of the form

$$SS = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

is minimised in order to calculate the parameters.

Where we believe the relationship is non-linear we can use a regression model of the form

$$Y_i = \alpha X_{1i}(1 + \beta X_{2i}) + \varepsilon_i.$$

This can be transformed into a form that is linear and then fitted using least squares. This is carried out by minimising squares and is described by

$$SS = \sum_{i=1}^{n} (y_i - \alpha(1 - e^{-\beta x_i}))^2.$$

Yield curve fitting techniques that use splines are often fitted using multiple regression methods.

Stochastic processes

This is perhaps the most difficult area of financial mathematics. Most references are also very technical and therefore difficult to access for the non-mathematician.

We begin with some definitions. A random process is usually referred to as a *stochastic* process. This is a collection of random variables $X(t)$ and the process may be either discrete or continuous. We write $\{X(t),\ t \in T\}$ and a sample $\{x(t),\ 0 \leq t \leq t_{\max}\}$ of the random process $\{X(t),\ t \geq 0\}$ is known as the *realisation* or *path* of the process.

A *Markov* process is one where the path is dependent on the present state of the process only, so that all historical data, including the path taken to arrive at the present state, is irrelevant. So in a Markov process, all data up to the present is contained in the present state. The dynamics of asset prices are frequently assumed to follow a Markov process, and in fact it represents a semi-strong form efficient market. It is written

$$Pr\{X(t) \leq y | X(u) = x(u),\ 0 \leq u \leq s\} = Pr\{X(t) \leq y | X(s) = x(s)\}$$

for $0 \leq s \leq t$.

A *Weiner process* or *Brownian motion* for $\{X(t),\ t \geq 0\}$ has the following properties:

$X(0) = 0$;
$\{X(t),\ t \geq 0\}$ has independent increments, so that $X(t+b) - X(t)$ and $X(t+2b) - X(t+b)$ are independent and follow the same distribution;
the variable $X(t)$ has the property $X(t) \sim N(0,\ t)$ for all $t > 0$;
$X(t) - X(s) \sim N(0,\ t - s)$ for $0 \leq s < t$.

Many interest-rate models assume that the movement of interest rates over time follows a Weiner process.

Stochastic calculus

The Weiner process is usually denoted with W although Z and z are also used. For a Weiner process $\{W(t),\ t \geq 0\}$ it can be shown that after an infinitesimal time interval Δt we have

$$W(t + \Delta t) - W(t) \sim N(0,\ \Delta t).$$

If we also have $U \sim N(0,\ 1)$ then we may write

$$W(t + \Delta t) - W(t) = \sqrt{\Delta t}U.$$

As the time interval decreases and approaches (but does not reach) 0, then the expression above may be written

$$dW(t) = \sqrt{\Delta t}U.$$

A Weiner process is not differentiable but a generalised Weiner process termed an Itô process is differentiable and is described in the form

$$dX(t) = a(t, X)dt + b(t, X)dW$$

where a is the drift and b the noise or volatility of the stochastic process.

Selected bibliography

Baxter, A., Rennie, D., *Financial Calculus*, Cambridge University Press, 1996.
Gujarati, D., *Basic Econometrics*, McGraw-Hill, 1997.
Hogg, R., Craig, A., *An Introduction to Mathematical Statistics*, 5th edition, Prentice Hall, 1995.
Kreyszig, E., *Advanced Engineering Mathematics*, Wiley, 1993.
Ross, S., *An Introduction to Probability Models*, 5th edition, Academic Press, 1997.
Ross, S., *An Introduction to Mathematical Finance*, Cambridge University Press, 1999.

1 Approaches to Trading and Hedging

It is not the intention of this book to suggest trading 'strategies' as such, or any particular approach to running a fixed-income market making or proprietary trading book. Rather, we will discuss certain approaches that have worked in the past and should, given the right circumstances, work again at some point in the future. It is the intention however, to focus on real-world application whilst maintaining analytical rigour. The term *trading* covers a wide range of activity. Market makers who are quoting two-way prices to market participants may be tasked with providing a customer service, building up retail and institutional volume, or they may be tasked with purely running the book at a profit and trying to maximise return on capital. The nature of the market that is traded will also impact on their approach. In a highly transparent and liquid market such as the US Treasury or the UK gilt market the price spreads are fairly narrow,[1] although increased demand has reduced this somewhat in both markets. However this means that opportunities for profitable trading as a result of mispricing of individual securities, whilst not completely extinct, are rare. It is much more common for traders in such markets to take a view on *relative value*

[1] In fact, in the late 1990s spreads in the gilt market were beginning to widen as excess demand over supply, particularly at the long end of the yield curve, drove down yields and reduced liquidity. In the Treasury market at the start of 2000 the yield curve had inverted, with the yield on the long bond at 6.05% in February that year, down over 70 basis points from the start of the year. The announcement by the US Treasury that it would buy back over $30 billion of debt in the year also led to increased demand at the long end, helping to depress yields. The volatility level in the market was at a two-year high at that time. These developments in the two markets have led to wider price quotes and lower liquidity. A sustained public sector deficit has many implications for the debt markets, if governments start to repay national debt and cease issuing securities. This is an important topic which is currently the subject of some debate. A significant reduction in government debt levels, while advantageous in many respects, will pose new problems. This is because government bonds play an important part in the financial systems of many countries. In the first instance, government bonds are used as the benchmark against which many other instruments are priced, such as corporate bonds. An illiquid market in government debt could have serious implications for the corporate bond markets, with investors possibly becoming reluctant to invest in corporate paper unless yield levels rise. Derivatives may also suffer from pricing problems, particularly bond futures contracts. In addition, while long-horizon institutional investors such as pension funds may find themselves short of investment products, many central banks and sovereign governments are big holders of securities such as US Treasuries, gilts and bunds. A shortage of supply in these instruments, particularly Treasuries, might have implications for all these investors unless an alternative instrument is made available. The continuing inverted yield curve in the UK, which dated from July 1997, and the inversion of the US curve in February 2000, is put down partly to shortage of long-dated government stock. The OECD, as reported in *The Economist* (12 February 2000) has suggested a policy whereby governments maintain a minimum level of gross public debt, with this minimum being an amount sufficient to maintain bond market liquidity. This may not be a practical solution for large economies however, especially that of the United States, but is certainly viable for other developed economies. The issue of alternative benchmarks is currently being researched by the author.

trades, such as the yield spreads between individual securities or the expected future shape of the yield curve. This is also called *spread trading*. A large volume of trading on derivatives exchanges is done for hedging purposes, but speculative trading is also prominent. Very often bond and interest-rate traders will punt using futures or options contracts, based on their view of market direction. Ironically, market makers who have a low level of customer business, perhaps because they are newcomers to the market, for historical reasons or because they do not have the appetite for risk that is required to service the high quality customers, tend to speculate on the futures exchanges to relieve tedium, often with unfortunate results.

Speculative trading is undertaken on the basis of the views of the trader, desk or head of the department. This view may be an 'in-house' view, for example the collective belief of the economics or research department, or the individual trader's view, which will be formulated as a result of *fundamental analysis* and *technical analysis*. The former is an assessment of macroeconomic and microeconomic factors affecting not just the specific bond market itself but the economy as a whole. Those running corporate debt desks will also concentrate heavily on individual sectors and corporations and their wider environment, because the credit spread, and what drives the credit spread, of corporate bonds is of course key to the performance of the bonds. Technical analysis or *charting* is a discipline in its own right, and has its adherents. It is based on the belief that over time the patterns displayed by a continuous time series of asset prices will repeat themselves. Therefore detecting patterns should give a reasonable expectation of how asset prices should behave in the future. Many traders use a combination of fundamental and technical analysis, although chartists often say that for technical analysis to work effectively, it must be the only method adopted by the trader. A review of technical analysis is presented in Chapter 63 of the author's book *The Bond and Money Markets*.

In this chapter we introduce some common methods and approaches, and some not so common, that might be employed on a fixed interest desk.

1.1 Futures trading

Trading with derivatives is often preferred, for both speculative or hedging purposes, to trading in the cash markets mainly because of the liquidity of the market and the ease and low cost of undertaking transactions. The essential features of futures trading are volatility and leverage. To establish a futures position on an exchange, the level of margin required is very low proportional to the notional value of the contracts traded. For speculative purposes traders often carry out open, that is uncovered trading, which is a directional bet on the market. So therefore if a trader believed that short-term sterling interest rates were going to fall, they could buy a short sterling contract on LIFFE. This may be held for under a day, in which case if the price rises the trader will gain, or for a longer period, depending on their view. The tick value of a short sterling contract is £10, so if they bought one lot at 92.75 (that is, 100 − 92.75 or 7.25%) and sold it at the end of the day for 98.85 they made a profit of £100 on their one lot, from which brokerage will be subtracted. The trade can be carried out with any futures contract; the same idea could be carried out with a cash market product or a FRA, but the liquidity, narrow price spread and the low cost of dealing make such a trade easier on a futures exchange. It is much more interesting however to carry out a spread trade on the difference between the rates of two different contracts. Consider Figures 1.1

and 1.2 which relate to the prices for the LIFFE short-sterling futures contract on 22 March 1999. The specification for this contract is summarised in Chapter 35 of Choudhry (2001).

Most futures exchanges use the designatory letters H, M, U and Z to refer to the contract months for March, June, September and December. So the June 1999 contract is denoted by 'M99'. From Chapter 2 we know that forward rates can be calculated for any term, starting on any date. In Figure 1.1 we see the future prices on that day, and the interest rate that the prices imply. The 'stub' is the term for the interest rate from today to the expiry of the first futures contract, which is called the *front month* contract (in this case the front month contract is the June 1999 contract). Figure 1.2 lists the forward rates from the spot date to six months, one year and so on. It is possible to trade a strip of contracts to replicate any term, out to the maximum maturity of the contract. This can be done for hedging or speculative purposes. Note from Figure 1.1 that there is a spread between the cash curve and the futures curve. A trader can take positions on cash against futures, but it is easier to transact only on the futures exchange.

Short-term money market interest rates often behave independently of the yield curve as a whole. A money markets trader may be aware of cash market trends, for example an increased frequency of borrowing at a certain point of the curve, as well as other market intelligence that suggests that one point of the curve will rise or fall relative to others. One way to exploit this view is to run a position in a cash instrument such as CD against a futures contract, which is a *basis spread* trade.

Date	**22/03/1999**	**LIFFE SHORT STERLING CONTRACT**								
Term	1w	1m	2m	3m	4m	5m	6m	9m	1y	
Libor	5.561	5.439	5.407	5.359	5.344	5.321	5.304	5.295	5.300	
Futures	M99	U99	Z99	H00	M00	U00	Z00	H01	M01	U01
Price	94.88	94.97	94.75	94.85	94.77	94.68	94.56	94.58	94.56	94.57
Rate %	5.12	5.03	5.25	5.15	5.23	5.32	5.44	5.42	5.44	5.43
Expiry	16-Jun	15-Sep	15-Dec	15-Mar	21-Jun	20-Sep	20-Dec	21-Mar	20-Jun	19-Sep
Days	86	177	268	359	457	548	639	730	821	912
Yield curve %				1y				2y		2.5y
Cash	5.367	5.307	5.295	5.300	5.243	5.252	5.262	5.270	5.287	5.303
Futures	5.367	5.273	5.234	5.289	5.214	5.223	5.242	5.271	5.291	5.310

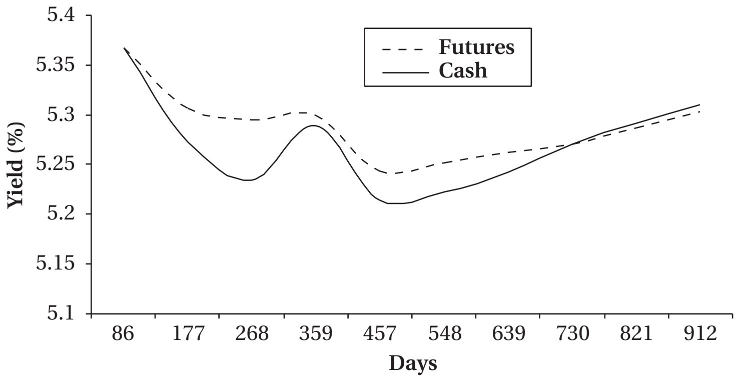


Figure 1.1: LIFFE short-sterling contract analysis, 22 March 1999.

Date	**22/03/1999**		**LIFFE SHORT STERLING CONTRACT**							
Date	**Days**	**Contract**	**Price**	**Rate**	**6m-fwd**	**1y-fwd**	**1.5y-fwd**	**2y-fwd**	**3y-fwd**	**4y-fwd**
Spot	86	Stub	94.6317	5.3683	5.268	5.292	5.224	5.271	5.337	5.356
16-Jun-99	91	M99	94.88	5.1200	5.107	5.239	5.216	5.278	5.341	5.353
15-Sep-99	91	U99	94.97	5.0300	5.173	5.267	5.269	5.318	5.363	
15-Dec-99	91	Z99	94.75	5.2500	5.233	5.341	5.334	5.369	5.390	
15-Mar-00	98	H00	94.85	5.1500	5.223	5.390	5.366	5.392	5.399	
21-Jun-00	91	M00	94.77	5.2300	5.310	5.461	5.416	5.428	5.415	
20-Sep-00	91	U00	94.68	5.3200	5.416	5.515	5.452	5.447		
20-Dec-00	91	Z00	94.56	5.4400	5.467	5.544	5.467	5.452		
21-Mar-01	91	H01	94.58	5.4200	5.467	5.544	5.457	5.440		
20-Jun-01	91	M01	94.56	5.4400	5.472	5.541	5.447	5.428		
19-Sep-01	91	U01	94.57	5.4300	5.472	5.526	5.431			

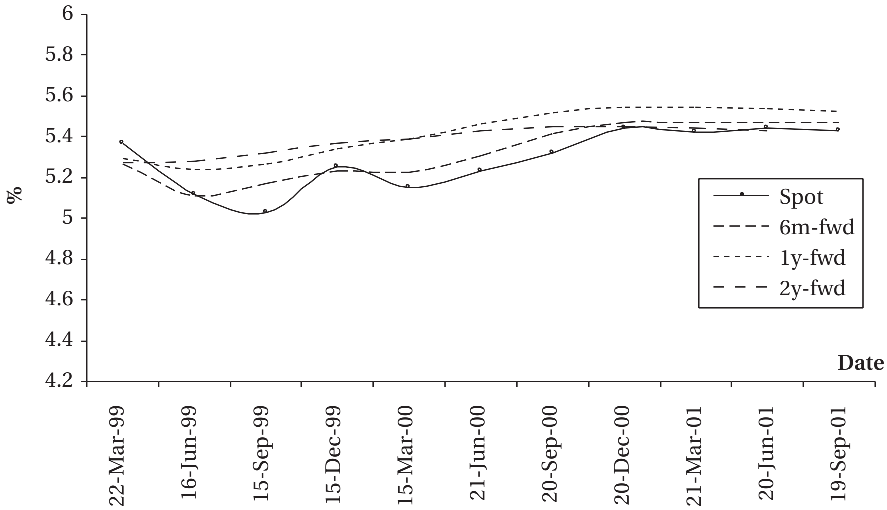


Figure 1.2: LIFFE short-sterling forward rates analysis, 22 March 1999.

However the best way to trade on this view is to carry out a spread trade, shorting one contract against a long position in another trade. Consider Figure 1.1; if we feel that three-month interest rates in June 2000 will be lower than where they are implied by the futures price today, but that September 2000 rates will be higher, we will buy the M00 contract and short the U00 contract. This is not a market directional trade, rather a view on the relative spread between two contracts. The trade must be carried out in equal weights, for example 100 lots of the June against 100 lots of the September.[2] If the rates do move in the

[2] The author would particularly like to thank Peter Matthews, now with ABN Amro Securities (UK) Ltd, and Ed Hardman who was in the short sterling booth at GNI on LIFFE, with whom I've sadly lost contact, for information on this type of trading back in 1994.

direction that the trader expects, the trade will generate a profit. There are similar possibilities available from an analysis of Figure 1.2, depending on our view of forward interest rates.

Spread trading carries a lower margin requirement than open position trading, because there is no directional risk in the trade. It is also possible to arbitrage between contracts on different exchanges. If the trade is short the near contract and long the far contract, so the opposite of our example, this is known as *buying the spread* and the trader believes the spread will widen. The opposite is *shorting the spread* and is undertaken when the trader believes the spread will narrow. Note that the difference between the two price levels is not limitless, because the theoretical price of a futures contract provides an upper limit to the size of the spread or the basis. The spread or the basis cannot exceed the cost of carry, that is the net cost of buying the cash security today and then delivering it into the futures market on the contract expiry. The same principle applies to short-dated interest-rate contracts; the net cost is the difference between the interest cost of borrowing funds to buy the 'security' and the income accruing on the security while it is held before delivery. The two associated costs for a short-sterling spread trade are the notional borrowing and lending rates from having bought one and sold another contract. If the trader believes that the cost of carry will decrease they could sell the spread to exercise this view.

The trader may have a longer time horizon and trade the spread between the short-term interest-rate contract and the long bond future. This is usually carried out only by proprietary traders, because it is unlikely that one person would be trading both three-month and 10-year (or 20-year, depending on the contract specification) interest rates. A common example of such a spread trade is a yield curve trade. If a trader believes that the sterling yield curve will steepen or flatten between the three-month and the 10-year terms, they can buy or sell the spread by using the LIFFE short-sterling contract and the long gilt contract. To be first-order risk neutral however the trade must be duration-weighted, as one short-sterling contract is not equivalent to one gilt contract. The tick value of both contracts is £10, although the gilt contract represents £100,000 of a notional gilt and the short-sterling contract represents a £500,000 time deposit. We use (1.1) to calculate the hedge ratio, with £1000 being the value of a 1% change in the value of both contracts.

$$h = \frac{(100 \times tick) \times P_b^f \times D}{(100 \times tick) \times P_{short\ ir}^f} \tag{1.1}$$

where

$tick$	is the tick value of the contract
D	is the duration of the bond represented by the long bond contract
P_b^f	is the price of the bond futures contract
$P_{short\ ir}^f$	is the price of the short-term deposit contract.

The notional maturity of a long bond contract is always given in terms of a spread, for example for the long gilt it is $8\frac{3}{4}$ –13 years. Therefore in practice one would use the duration of the cheapest-to-deliver bond.

A *butterfly spread* is a spread trade that involves three contracts, with the two spreads between all three contracts being traded. This is carried out when the middle contract appears to be mispriced relative to the two contracts either side of it. The trader may believe that one or both of the outer contracts will move in relation to the middle contract; if the belief is that only one of these two will shift relative to the middle contract, then a butterfly will be put on if the trader is not sure which of these will adjust. For example, consider Figure 1.1 again. The prices of the front three contracts are 94.88, 94.97 and 94.75. A trader may feel that the September contract is too low, and has a spread of +9 basis points to the June contract, and −22 basis points to the December contract. The trader feels that the September contract will rise, but will that be because June and December prices fall or because the September price will rise? Instead of having to answer this question, all the trader need believe is that the Jun–Sep spread will widen and the Sep–Dec spread will narrow. To put this view into effect, the trader puts on a butterfly spread, which is equal to the Sep–Dec spread minus the Jun–Sep spread, which they expect to narrow. Therefore the trader buys the Jun–Sep spread and sells the Sep–Dec spread, which is also known as *selling the butterfly spread*.

1.2 Yield curves and relative value

Bond market participants take a keen interest in the yield curve, both cash and zero-coupon (spot) yield curves. In markets where an active zero-coupon bond market exists, much analysis is undertaken into the relative spreads between derived and actual zero-coupon yields. In this section we review some of the yield curve analysis used in the market.

1.2.1 *The determinants of government bond yields*

Market makers in government bond markets will analyse various factors in the market in deciding how to run their book. Customer business apart, decisions to purchase or sell securities will be a function of their views on:

- market direction itself, that is the direction in which short-term and long-term interest rates are headed;
- which maturity point along the entire term structure offers the best value;
- which specific issue within a particular maturity point offers the best value.

All three areas are related but will react differently to certain pieces of information. A report on the projected size of the government's budget deficit for example, will not have much effect on two-year bond yields, whereas if the expectations came as a surprise to the market it could have an adverse effect on long-bond yields. The starting point for analysis is of course the yield curve, both the traditional coupon curve plotted against duration and the zero-coupon curve. Figure 1.3 illustrates the traditional yield curve for gilts in October 1999.

For a first-level analysis, many market practitioners will go no further than Figure 1.3. An investor who had no particular view on the future shape of the yield curve or the level of interest rates may well adopt a neutral outlook and hold bonds that have a duration that matches their investment horizon. If they believed interest rates were likely to remain stable for a time, they might hold bonds with a longer duration in a positive sloping yield curve environment, and pick up additional yield but with higher interest-rate risk. Once the decision has been made on which part of the yield curve to invest in or switch into, the investor must decide on the specific securities to hold, which then brings us on to relative

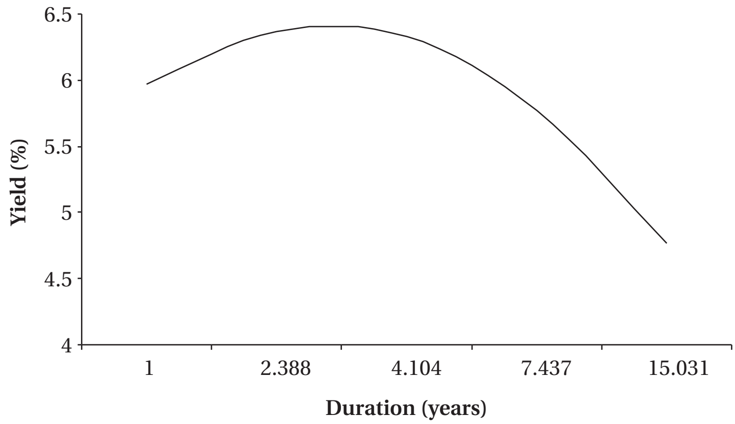


Figure 1.3: Yield and duration of gilts, 21 October 1999.

value analysis. For this the investor will analyse specific sectors of the curve, looking at individual stocks. This is sometimes called looking at the 'local' part of the curve.

An assessment of a local part of the yield curve will include looking at other features of individual stocks in addition to their duration. This recognises that the yield of a specific bond is not only a function of its duration, and that two bonds with near-identical duration can have different yields. The other determinants of yield are liquidity of the bond and its coupon. To illustrate the effect of coupon on yield consider Table 1.1. This shows that, where the duration of a bond is held roughly constant, a change in coupon of a bond can have a significant effect on the bond's yield.

In the case of the long bond, an investor could under this scenario both shorten duration and pick up yield, which is not the first thing that an investor might expect. However an anomaly of the markets is that, liquidity issues aside, the market does not generally like high coupon bonds, so they usually trade cheap to the curve.

The other factors affecting yield are supply and demand, and liquidity. A shortage of supply of stock at a particular point in the curve will have the effect of depressing yields at that point. A reducing public sector deficit is the main reason why such a supply shortage might exist. In addition as interest rates decline say ahead of or during a recession, the stock of high coupon bonds increases, as the newer bonds are issued at lower levels, and these 'outdated' issues can end up trading at a higher yield. Demand factors are driven primarily by the investor's views of the country's economic prospects, but also by government

Coupon	Maturity	Duration	Yield
8%	20-Feb-02	1.927	5.75%
12%	5-Feb-02	1.911	5.80%
10%	20-Jun-10	7.134	4.95%
6%	1-Jul-10	7.867	4.77%

Table 1.1: Duration and yield comparisons for bonds in a hypothetical inverted curve environment.

legislation, for example the Minimum Funding Requirement in the UK compels pension funds to hold a set minimum amount of their funds in long-dated gilts, which has the effect of permanently keeping demand high.[3]

Liquidity often results in one bond having a higher yield than another, despite both having similar durations. Institutional investors prefer to hold the benchmark bond, which is the current two-year, five-year, ten-year or thirty-year bond and this depresses the yield on the benchmark bond. A bond that is liquid also has a higher demand, thus a lower yield, because it is easier to convert into cash if required. This can be demonstrated by valuing the cash flows on a six-month bond with the rates obtainable in the Treasury bill market. We could value the six-month cash flows at the six-month bill rate. The lowest obtainable yield in virtually every market[4] is the T-bill yield, therefore valuing a six-month bond at the T-bill rate will produce a discrepancy between the observed price of the bond and its theoretical price implied by the T-bill rate; as the observed price will be lower. The reason for this is simple: because the T-bill is more readily realisable into cash at any time, it trades at a lower yield than the bond, even though the cash flows fall on the same day.

We have therefore determined that a bond's coupon and liquidity level, as well as its duration, will affect the yield at which it trades. These factors can be used in conjunction with other areas of analysis, which we look at next, when deciding which bonds carry relative value over others.

1.2.2 *Characterising the complete term structure*

As many readers would have gathered, the yield versus duration curve illustrated in Figure 1.3 is an ineffective technique with which to analyse the market.

This is because it does not highlight any characteristics of the yield curve other than its general shape; this does not assist in the making of trading decisions. To facilitate a more complete picture, we might wish to employ the technique described here. Figure 1.4 shows the bond par yield curve[5] and T-bill yield curve for gilts in October 1999. Figure 1.5 shows the difference between the yield on a bond with a coupon that is 100 basis points below the par yield level, and the yield on a par bond. The other curve in Figure 1.5 shows the level for a bond with a coupon that is 100 basis points above the par yield. These two curves show the 'low coupon' and 'high coupon' yield spreads. Using the two figures together, an investor can see the impact of coupons, the shape of the curve and the effect of yield on different maturity points of the curve.

1.2.3 *Identifying relative value in government bonds*

Constructing a zero-coupon yield curve provides the framework within which a market participant can analyse individual securities. In a government bond market, there is no credit risk consideration (unless it is an emerging market government market), and

[3] The requirements of the MFR were removed during 2002 and the UK gilt yield curve exhibited a conventional positive-sloping shape shortly afterwards.

[4] The author is not aware of any market where there is a yield lower than its shortest-maturity T-bill yield, but that does not mean such a market doesn't exist!

[5] See Chapter 2 in the companion volume to this book in the Fixed Income Markets Library, *Corporate Bonds and Structured Financial Products*, for a discussion of the par yield curve.

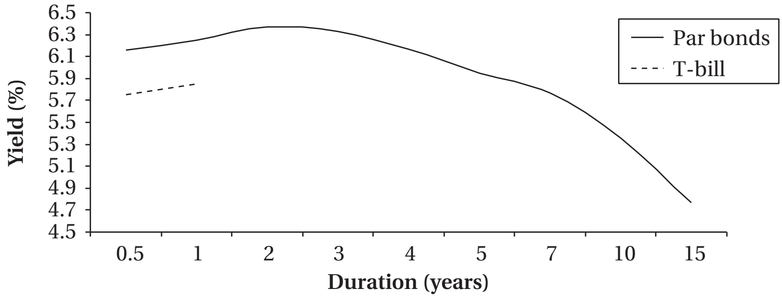


Figure 1.4: T-bill and par yield curve, October 1999.

therefore no credit spreads to consider. There are a number of factors that can be assessed in an attempt to identify relative value.

The objective of much of the analysis that occurs in bond markets is to identify value, and identifying which individual securities should be purchased and which sold. At the overview level, this identification is a function of whether one thinks interest rates are going to rise or fall. At the local level though, the analysis is more concerned with a specific sector of the yield curve, whether this will flatten or steepen, whether bonds of similar duration are trading at enough of a spread to warrant switching from one into another. The difference in these approaches is one of identifying which stocks have absolute value, and which have relative value. A trade decision based on the expected direction of interest rates is based on assessing absolute value, whether interest rates themselves are too low or too high. Yield curve analysis is more a matter of assessing relative value. On (very!) rare occasions, this process is fairly straightforward, for example if the three-year bond is trading at 5.75% when two-year yields are 5.70% and four-year yields are at 6.15%, the three-year would appear to be overpriced. However this is not really a real-life situation. Instead, a trader might find himself assessing the relative value of the three-year bond compared to much shorter- or longer-dated instruments. That said, there is considerable difference between comparing a short-dated bond to other short-term securities and comparing say, the two-year bond to

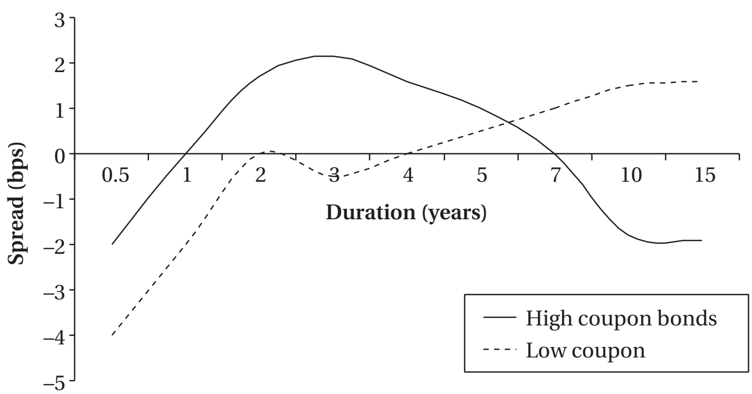


Figure 1.5: Structure of bond yields, October 1999.

the thirty-year bond. Although it looks like it on paper, the space along the x-axis should not be taken to imply that the smooth link between one-year and five-year bonds is repeated from the five-year out to the thirty-year bonds. It is also common for the very short-dated sector of the yield curve to behave independently of the long end.

One method used to identify relative value is to quantify the coupon effect on the yields of bonds. The relationship between yield and coupon is given by (1.2):

$$rm = rm_P + c \cdot max(C_{PD} - rm_P,\ 0) + d \cdot min(C_{PD} - rm_P,\ 0) \tag{1.2}$$

where

rm is the yield on the bond being analysed

rm_P is the yield on a par bond of specified duration

C_{PD} is the coupon on an arbitrary bond of similar duration to the part bond

and c and d are coefficients. The coefficient c reflects the effect of a high coupon on the yield of a bond. If we consider a case where the coupon rate exceeds the yield on the similar-duration par bond ($C_{PD} > rm_P$), (1.2) reduces to (1.3):

$$rm = rm_P + c \cdot (C_{PD} - rm_P). \tag{1.3}$$

Equation (1.3) specifies the spread between the yield on a high coupon bond and the yield on a par bond as a linear function of the spread between the first bond's coupon and the yield and coupon of the par bond. In reality this relationship may not be purely linear; for instance the yield spread may widen at a decreasing rate for higher coupon differences. Therefore (1.3) is an approximation of the effect of a high coupon on yield where the approximation is more appropriate for bonds trading close to par. The same analysis can be applied to bonds with coupons lower than the same-duration par bond.

The value of a bond may be measured against comparable securities or against the par or zero-coupon yield curve. In certain instances the first measure may be more appropriate when for instance, a low coupon bond is priced expensive to the curve itself but fair compared to other low coupon bonds. In that case the overpricing indicated by the par yield curve may not represent unusual value, rather a valuation phenomenon that was shared by all low coupon bonds. Having examined the local structure of a yield curve, the analysis can be extended to the comparative valuation of a group of similar bonds. This is an important part of the analysis, because it is particularly informative to know the cheapness or dearness of a single stock compared to the whole yield curve, which might be somewhat abstract. Instead we would seek to identify two or more bonds, one of which was cheap and the other dear, so that we might carry out an outright switch between the two, or put on a spread trade between them. Using the technique we can identify excess positive or negative yield spread for all the bonds in the term structure. This has been carried out for our five gilts, together with other less liquid issues as at October 1999 and the results are summarised in Table 1.2.

From the table as we might expect the benchmark securities are all expensive to the par curve, and the less liquid bonds are cheap. Note that the 6.25% 2010 appears cheap to the curve, but the 5.75% 2009 offers a yield pick-up for what is a shorter-duration stock; this is a curious anomaly and one that had disappeared a few days later.[6]

---

[6] In other words, we've missed the opportunity! This analysis used mid-prices, which would not be available in practice.

Coupon	Maturity	Duration	Yield %	Excess yield spread (bp)
8%	07/12/2000	1.072	5.972	−1.55
10%	26/02/2001	1.2601	6.051	4.5
7%	07/06/2002	2.388	6.367	−1.8
5%	07/06/2004	4.104	6.327	−3.8
6.75%	26/11/2004	4.233	6.351	2.7
5.75%	07/12/2009	7.437	5.77	−4.7
6.25%	25/11/2010	7.957	5.72	1.08
6%	07/12/2028	15.031	4.77	−8.7

Table 1.2: Yields and excess yield spreads for selected gilts, 22 October 1999.

What this section has introduced is the concept of relative value for individual securities, and how the simple duration/yield analysis can be extended to assess other determinants of a bond's yield. We now look at the issues involved in putting on a spread trade.

1.3 Yield spread trades

In the earlier section on futures trading, we introduced the concept of spread trading, which are not market directional trades but rather the expression of a viewpoint on the shape of a yield curve, or more specifically the spread between two particular points on the yield curve. Generally there is no analytical relationship between changes in a specific yield spread and changes in the general level of interest rates. That is to say, the yield curve can flatten when rates are both falling or rising, and equally may steepen under either scenario as well. The key element of any spread trade is that it is structured so that a profit (or any loss) is made only as a result of a change in the spread, and not due to any change in overall yield levels. That is, spread trading eliminates market directional or first-order market risk.

1.3.1 *Bond spread weighting*

Table 1.3 shows data for our selection of gilts but with additional information on the basis point value (BPV) for each point. This is also known as the 'dollar value of a basis point' or DV01.

If a trader believed that the yield curve was going to flatten, but had no particular strong feeling about whether this flattening would occur in an environment of falling or rising interest rates, and thought that the flattening would be most pronounced in the two-year versus ten-year spread, they could put on a spread consisting of a short position in the two-year and a long position in the ten-year. This spread must be duration-weighted to eliminate first-order risk. At this stage we must point out, and it is important to be aware of, the fact that basis point values, which are used to weight the trade, are based on modified duration measures. This measure is an approximation, and will be inaccurate for large changes in yield. Therefore the trader must monitor the spread to ensure that the weights are not going out of line, especially in a volatile market environment.

Coupon	Maturity	Duration	Yield %	Price	BPV
8%	07/12/2000	1.072	5.972	102.17	0.01095
10%	26/02/2001	1.2601	6.051	105.01	0.01880
7%	07/06/2002	2.388	6.367	101.5	0.02410
5%	07/06/2004	4.104	6.327	94.74	0.03835
6.75%	26/11/2004	4.233	6.351	101.71	0.03980
5.75%	07/12/2009	7.437	5.77	99.84	0.07584
6.25%	25/11/2010	7.957	5.72	104.3	0.07526
6%	07/12/2028	15.031	4.77	119.25	0.17834

Table 1.3: Bond basis point value, 22 October 1999.

To weight the spread, we use the ratios of the BPVs of each bond to decide on how much to trade. In our example, assume the trader wants to purchase £10 million of the ten-year. In that case he must sell ((0.07584/0.02410) $\times$ 10,000,000) or £31,468,880 of the two-year bond. It is also possible to weight a trade using the bonds' duration values, but this is rare. It is common practice to use the BPV.

The payoff from the trade will depend on what happens to the two-year versus ten-year spread. If the yields on both bonds move by the same amount, there will be no profit generated, although there will be a funding consideration. If the spread does indeed narrow, the trade will generate profit. Note that disciplined trading calls for both an expected target spread as well as a fixed time horizon. So for example, the current spread is 59.7 basis points; the trader may decide to take the profit if the spread narrows to 50 basis points, with a three-week horizon. If at the end of three weeks the spread has not reached the target, the trader should unwind the position anyway, because that was their original target. On the other hand what if the spread has narrowed to 48 basis points after one week and looks like narrowing further – what should the trader do? Again, disciplined trading suggests the profit should be taken. If contrary to expectations the spread starts to widen, if it reaches 64.5 basis points the trade should be unwound, this 'stop-loss' being at the half-way point of the original profit target.

The financing of the trade in the repo markets is an important aspect of the trade, and will set the trade's break-even level. If the bond being shorted (in our example, the two-year bond) is *special*, this will have an adverse impact on the financing of the trade. The repo considerations are reviewed in Choudhry (2002).

1.3.2 *Types of bond spreads*

A bond spread has two fundamental characteristics; in theory there should be no P/L effect due to a general change in interest rates, and any P/L should only occur as a result of a change in the specific spread being traded. Most bond spread trades are yield curve trades where a view is taken on whether a particular spread will widen or narrow. Therefore it is important to be able to identify which sectors of the curve to sell. Assuming that a trader is able to transact business along any part of the yield curve, there are a number of factors to consider. In the first instance, the historic spread between the two sectors of the curve. To illustrate in simplistic fashion, if the 2–10 year spread has been between 40 and 50 basis

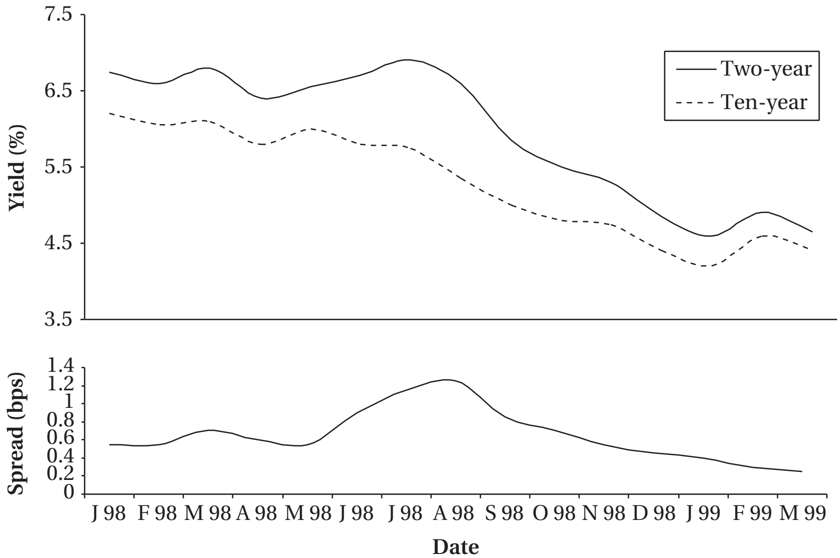


Figure 1.6: 2-year and 10-year spread, UK gilt market March 1999. Source: Bloomberg.

points over the last six months but very recently has narrowed to less than 35 basis points, this may indicate imminent widening. Other factors to consider are demand and liquidity for individual stocks relative to others, and any market intelligence that the trader gleans. If there has been considerable customer interest on certain stocks relative to others, because investors themselves are switching out of certain stocks and into others, this may indicate a possible yield curve play. It is a matter of individual judgement.

An historical analysis requires that the trader identify some part of the yield curve within which he expects to observe a flattening or steepening. It is of course entirely possible that one segment of the curve will flatten while another segment is steepening, in fact this phenomenon is quite common. This reflects the fact that different segments respond to news and other occurrences in different ways.

A more exotic type of yield curve spread is a *curvature* trade. Consider for example a trader who believes that three-year bonds will outperform on a relative basis, both two-year and five-year bonds. That is, he believes that the two-year/three-year spread will narrow relative to the three-year/five-year spread, in other words that the curvature of the yield curve will decrease. This is also known as a *butterfly/barbell* trade. In our example the trader will buy the three-year bond, against short sales of both the two-year and the five-year bonds. All positions are duration-weighted. The principle is exactly the same as the butterfly trade we described in the previous section on futures trading.

1.4 Hedging bond positions

Hedging is a straightforward concept to understand or describe, however it is very important that it is undertaken as accurately as possible. Therefore the calculation of a hedge is critical. A hedge is a position in a cash or off-balance sheet instrument that removes the market risk exposure of another position. For example a long position in 10-year bonds can be hedged with a short position in 20-year bonds, or with futures contracts. That is the straightforward part; the calculation of the exact amount of the hedge is where complexities can arise. In this section we review the basic concepts of hedging, and a case study at the end illustrates some of the factors that must be considered.

1.4.1 *Simple hedging approach*

The hedge calculation that first presents itself is the duration-weighted approach. From the sample of gilts in Table 1.3, it is possible to calculate the amount of one bond required to hedge an amount of any other bond, using the ratio of the BPVs. This approach is very common in the market; however it suffers from two basic flaws that hinder its effectiveness. First, the approach assumes implicitly comparable volatility of yields on the two bonds, and secondly it also assumes that yield changes on the two bonds are highly correlated. Where one or both of these factors do not apply, the effectiveness of the hedge will be compromised.

The assumption of comparable volatility becomes increasingly unrealistic the more the bonds differ in terms of market risk and market behaviour. Consider a long position in two-year bonds hedged with a short-position in five-year bonds. Using the bonds from Table 1.3, if we had a position of £1 million of the two-year, we would short £628,422 of the five-year. Even if we imagine that yields between the two bonds are perfectly correlated, it may well be that the amount of yield change is different because the bonds have different volatilities. For example if the yield on the five-year bond changes only by half the amount that the two-year does, if there was a 5 basis point rise in the two-year, the five-year would have risen only by 2.5 basis points. This would indicate that the yield volatility of the two-year bond was twice that of the five-year bond. This suggests that a hedge calculation that matched nominal amounts, due to BPV, on the basis of an equal change in yield for both bonds would be incorrect. In our illustration, the short position in the five-year bond would be effectively hedging only half of the risk exposure of the two-year position.

The implicit assumption of perfectly correlated yield changes can also lead to inaccuracy. Across the whole term structure, it is not always the case that bond yields are even positively correlated all the time (although most of the time there will be a close positive correlation). Therefore, using our illustration again, imagine that the two-year and the five-year bonds possess identical yield volatilities, but that changes in their yields are uncorrelated. This means that knowing that the yield on the two-year bond rose or fell by one basis point does not tell us anything about the change in the yield on the five-year bond. If yield changes between the two bonds are indeed uncorrelated, this means that the five-year bonds cannot be used to hedge two-year bonds, at least not with accuracy.

1.4.2 *Hedge analysis*

From the foregoing we note that there are at least three factors that will impact the effectiveness of a bond hedge; these are the basis point value, the yield volatility of each bond and the correlation between changes in the two yields of a pair of bonds. Considering volatilities and correlations, Table 1.4 shows the standard deviations and correlations of weekly yield changes for a set of gilts during the nine months to October 1999. The standard deviation of weekly yield changes was in fact highest for the short-date paper, and actually

	Segment 2-year	3-year	5-year	10-year	20-year	30-year
Volatility (bp)	**19.3**	**19.5**	**20.2**	**20.0**	**20.1**	**20.3**
Correlation						
2-year	1.000	0.973	0.949	0.919	0.887	0.879
3-year	0.973	1.000	0.961	0.935	0.901	0.889
5-year	0.949	0.961	1.000	0.968	0.951	0.945
10-year	0.919	0.935	0.968	1.000	0.981	0.983
20-year	0.887	0.901	0.951	0.981	1.000	0.987
30-year	0.879	0.889	0.945	0.983	0.987	1.000

Table 1.4: Yield volatility and correlations, selected gilts October 1999.

declined for longer-dated paper. From the table we also note that changes in yield were imperfectly correlated. We expect correlations to be highest for bonds in the same segments of the yield curve, and to decline between bonds that are in different segments. This is not surprising, and indeed two-year bond yields are more positively correlated with five-year bonds and less so with 30-year bonds.

We can use the standard relationship for correlations and the effect of correlation to adjust a hedge. Consider two bonds with nominal values M_1 and M_2; if the yields on these two bonds change by Δr_1 and Δr_2 the net value of the change in position is given by:

$$\Delta PV = M_1 BPV_1 \Delta r_1 + M_2 BPV_2 \Delta r_2. \tag{1.4}$$

The uncertainty of the change in the net value of a two-bond position is dependent on the nominal values, the volatility of each bond and the correlation between these yield changes. Therefore for a two-bond position we set the standard deviation of the change in the position as (1.5):

$$\sigma_{pos} = \sqrt{M_1^2 BPV_1^2 \sigma_1^2 + M_2^2 BPV_2^2 \sigma_2^2 + 2M_1 M_2 BPV_1 BPV_2 \sigma_1 \sigma_2 \rho} \tag{1.5}$$

where ρ is the correlation between the yield volatilities of bonds 1 and 2. We can rearrange (1.5) to set the optimum hedge value for any bond using (1.6):

$$M_2 = -\frac{\rho BPV_1 \sigma_1}{BPV_2 \sigma_2} M_1 \tag{1.6}$$

so that M_2 is the nominal value of any bond used as a hedge given any nominal value M_1 of the first bond, and using each bond's volatility and the correlation. The derivation of (1.6) is given in Appendix 1.1. A lower correlation leads to a smaller hedge position, because where yield changes are not closely related, this implies greater independence between yield changes of the two bonds. In a scenario where the standard deviation of two bonds is identical, and the correlation between yield changes is 1, (1.6) reduces to:

$$M_2 = \frac{BPV_1}{BPV_2} M_1 \tag{1.7}$$

which is the traditional hedge calculation based solely on basis point values.

Case study:

Hedging a portfolio of Eurobonds with US Treasuries

Consider a portfolio of value $10,000,000 composed of the following US$ Eurobonds:
7% 2000
7.75% 2002
9% 2004
9% 2006

A trader is concerned that yields will rise over the next 48 hours, and decides to construct a short position of $100 million of US. Treasuries that will hedge the portfolio of US$ Eurobonds against the expected rise in yields. To determine the accuracy of hedge the trader will compare the change in value of the US$ Eurobond portfolio to that of the short position of the US Treasury hedge. All the possible hedge bonds under consideration are given below.

Bond	Dirty price	Duration	Convexity
T7% 1999	100.375512	1.852	4.165
T7% 2000	99.994565	2.623	7.987
T7% 2000	100.71875	3.089	10.917
T7% 2002	99.668545	4.176	19.823
T7% 2004	100.262295	5.284	32.634
T8% 2007	99.315217	6.905	57.913
T9¾% 1999	110.011395	8.168	90.636
T7½% 1999	90.654891	10.895	178.898
T8¾% 1999	106.116885	10.794	186.779
T8% 1999	97.769361	11.251	199.657

The Eurobond portfolio has the following initial values:

Value	Duration	Convexity
$10,000,000	4.550	27.218

The duration and market value can be matched analytically with any two bonds provided the duration of one of the bonds is less than that of the portfolio and the duration of the other bond is greater than that of the portfolio.

The trader might elect to do the following:

He can attempt to match both duration and convexity by constructing two portfolios with a duration of 4.550, one with a convexity greater than 27.218 and the other with a convexity less than 27.218. Assume these portfolios are called ‘A’ and ‘B’, respectively. Inspection of the table given above suggests the following bonds would be suitable components of these two portfolios:

A: T7% 2000 and T8% 2007
B: T7% 2002 and T7% 2004

Having identified two bonds with duration below and above 4.550 the trader then calculates the nominal value required for each bond using the simultaneous equation method giving an overall portfolio duration of 4.550.

Portfolio A:

Bond	Nominal required	Duration	Combined duration	Convexity
T7% 2002	6,644,185	4.176	4.550	24.151
T7% 2004	3,369,000	5.284		

Portfolio B:

As above, but with the following bonds:

Bond	Nominal required	Duration	Combined duration	Convexity
T7% 2000	6,127,300	3.089	4.550	28.911
T8% 2007	3,855,059	6.905		

There are now two portfolios with the same duration. Therefore, however the portfolios are combined, the duration will remain at 4.550. We now need to determine what amounts of each portfolio are required, such that the combined portfolio convexity is 27.218, matching the convexity of the Eurobond portfolio. The trader can use simple proportions to determine the amount of each portfolio which would be necessary to form a new portfolio of convexity 27.218. The exact amounts are:

A: 0.356 B: 0.644

Thus the composition of the hedging portfolio, which has a value of $10 million is:

T7% 2000	3,945,464 nominal
T7% 2002	2,365,484 nominal
T7% 2004	1,199,442 nominal
T8% 2007	2,482,828 nominal

Assume that the yield curve scenario two days later is a curvature twist around the 5-year maturity. The market value of the Eurobond portfolio improves by $75,984 to $10,075,984. The suggested combination of US Treasuries would mirror this gain as a loss – the suggested possible solution does result in a loss of $76,512 – a difference of $528.

Appendix

Appendix 1.1: Summary of derivation of optimum hedge equation

From equation (1.5) we know that the variance of a net change in the value of a two-bond portfolio is given by

$$\sigma_{pos}^2 = M_1^2 BPV_1^2 \sigma_1^2 + M_2^2 BPV_2^2 \sigma_2^2 + 2M_1 M_2 BPV_1 BPV_2 \sigma_1 \sigma_2 \rho. \tag{1.8}$$

Using the partial derivative of the variance σ^2 with respect to the nominal value of the second bond, we obtain

$$\frac{\partial \sigma^2}{\partial^2 M_2} = 2M_2 BPV_2^2 \sigma_2^2 + 2M_1 BPV_1 BPV_2 \sigma_1 \sigma_2 \rho. \tag{1.9}$$

If (1.8) is set to zero and solved for M_2 we obtain (1.10) which is the hedge quantity for the second bond

$$M_2 = -\frac{\rho BPV_1 \sigma_1}{BPV_2 \sigma_2} M_1. \tag{1.10}$$

Selected bibliography

Brennan, M., Schwartz, E., 'A continuous time approach to the pricing of bonds', *Journal of Banking and Finance* 3, 1979, p. 134 *ff*.

Brennan, M., Schwartz, E., 'Conditional predictions of bond prices and returns', *Journal of Finance* 35, 1980, p. 405 *ff*.

Campbell, J., Lo, A., MacKinley, A., *The Econometrics of Financial Markets*, Princeton University Press, 1997.

Choudhry, M., *Bond Market Securities*, FT Prentice Hall, 2001.

Choudhry, M., *The Repo Handbook*, Butterworth-Heinemann, 2002.

Fisher, L., Leibowitz, M., 'Effects of alternative anticipations of yield curve behaviour on the composition of immunized portfolios and on their target returns', in Kaufmann, G., *et al.* (eds.), *Innovations in Bond Portfolio Management*, Jai Press, 1983.

Neftci, S., *An Introduction to the Mathematics of Financial Derivatives*, 2nd edition, Academic Press, 2000.

The Economist, 'Out of Debt', *The Economist*, 12 February 2000.

2 Relative Value Trading using Government Bonds[1]

2.1 Introduction

Portfolio managers who do not wish to put on a naked directional position, but rather believe that the yield curve will change shape and flatten or widen between two selected points, put on relative value trades to reflect their view. Such trades involve simultaneous positions in bonds of different maturity. Other relative value trades may position high-coupon bonds against low-coupon bonds of the same maturity, as a tax-related transaction. These trades are concerned with the change in yield spread between two or more bonds rather than a change in absolute interest rate level. The key factor is that changes in spread are not conditional upon directional change in interest-rate levels; that is, yield spreads may narrow or widen whether interest rates themselves are rising or falling.

Typically, spread trades will be constructed as a long position in one bond against a short position in another bond. If it is set up correctly, the trade will only incur a profit or loss if there is a change in the shape of the yield curve. This is regarded as being *first-order risk neutral*, which means that there is no interest-rate risk in the event of change in the general level of market interest rates, provided the yield curve experiences essentially a parallel shift. In this chapter we examine some common yield spread trades.

2.1.1 *The determinants of yield*

The yield at which a fixed interest security is traded is market-determined. This market determination is a function of three factors: the term-to-maturity of the bond, the liquidity of the bond and its credit quality. Government securities such as gilts are default-free and so this factor drops out of the analysis. Under 'normal' circumstances the yield on a bond is higher the greater its maturity, this reflecting both the expectations hypothesis and liquidity preference theories. Intuitively we associate higher risk with longer-dated instruments, for which investors must be compensated in the form of higher yield. This higher risk reflects greater uncertainty with longer-dated bonds, both in terms of default and future inflation and interest rate levels. However, for a number of reasons the yield curve assumes an

[1] This chapter was presented by the author as an internal paper in July 1997 when he was working at Hambros Bank Limited. It previously appeared in Fabozzi (2002). The prices quoted are tick prices, fractions of 32nd, identical to US Treasury pricing. Gilts are now quoted as decimal prices.

inverted shape and long-dated yields become lower than short-dated ones.[2] Long-dated yields, generally, are expected to be less volatile over time compared to short-dated yields. This is mainly because incremental changes to economic circumstances or other technical considerations generally have an impact for only short periods of time, which affects the shorter end of the yield curve to a greater extent.

The liquidity of a bond also influences its yield level. The liquidity may be measured by the size of the bid–offer spread, the ease with which the stock may be transacted in size, and the impact of large-size bargains on the market. It is also measured by the extent of any *specialness* in its repo rate. Supply and demand for an individual stock, and the amount of stock available to trade, are the main drivers of liquidity.[3] The general rule is that there is a yield premium for transacting business in lower-liquidity bonds.

In the analysis that follows we assume satisfactory levels of liquidity, that is, it is straightforward to deal in large sizes without adversely moving the market.

2.1.2 *Spread trade risk weighting*

A relative value trade usually involves a long position set up against a short position in a bond of different maturity. The trade must be weighted so that the two positions are first-order neutral, which means the risk exposure of each position nets out when considered as a single trade, but only with respect to a general change in interest rate levels. If there is a change in yield spread, a profit or loss will be generated.

A common approach to weighting spread trades is to use the *basis point value* (BPV) of each bond.[4] Figure 2.1 shows price and yield data for a set of benchmark gilts for value date 17 June 1997.[5] The BPV for each bond is also shown, per £100 of stock. For the purposes of this discussion we quote mid-prices only and assume that the investor is able to trade at these prices. The yield curve at that date is shown in Figure 2.2.

The yield spread history between these two stocks over the previous three months and up to yesterday's closing yields is shown in Figure 2.3. An investor believes that the yield curve will flatten between the two-year and 10-year sectors of the curve and that the spread between the 6% 1999 and the 7.25% 2007 will narrow further from its present value of 0.299%.

To reflect this view the investor buys the 10-year bond and sells short the two-year bond, in amounts that leave the trade first-order risk neutral. If we assume the investor buys £1 million nominal of the 7.25% 2007 gilt, this represents an exposure of £1230.04 loss (profit) if there is a 1 basis point increase (decrease) in yields. Therefore, the nominal amount of the short position in the 6% 1999 gilt must equate this risk exposure. The BPV per £1 million nominal of the two-year bond is £166.42, which means that the investor must sell

---

2 For a summary of term structure theories see Chapter 2 of the companion volume to this book in the Fixed Income Markets Library, *Corporate Bonds and Structured Financial Products*.

3 The amount of stock issued and the amount of stock available to trade are not the same thing. If a large amount of a particular issue has been locked away by institutional investors, this may impede liquidity. However, the existence of a large amount at least means that some of the paper may be made available for lending in the stock loan and repo markets. A small issue size is a good indicator of low liquidity.

4 This is also known as *dollar value of a basis point* (DVBP or DV01) or *present value of a basis point* (PVBP).

5 Gilts settle on a $T + 1$ basis.

Term	Bond	Price	Accrued	Dirty price	Yield %	Modified duration	BPV	Per £1m nominal
2-yr	6% 10/8/1999	98-17	127	100.62	6.753	1.689	0.016642	166.42
5-yr	7% 7/6/2002	100-10	10	100.50	6.922	3.999	0.040115	401.15
10-yr	7.25% 7/12/2007	101-14	10	101.64	7.052	6.911	0.070103	701.03
25-yr	8% 7/6/2021	110-01	10	110.25	7.120	11.179	0.123004	1230.04

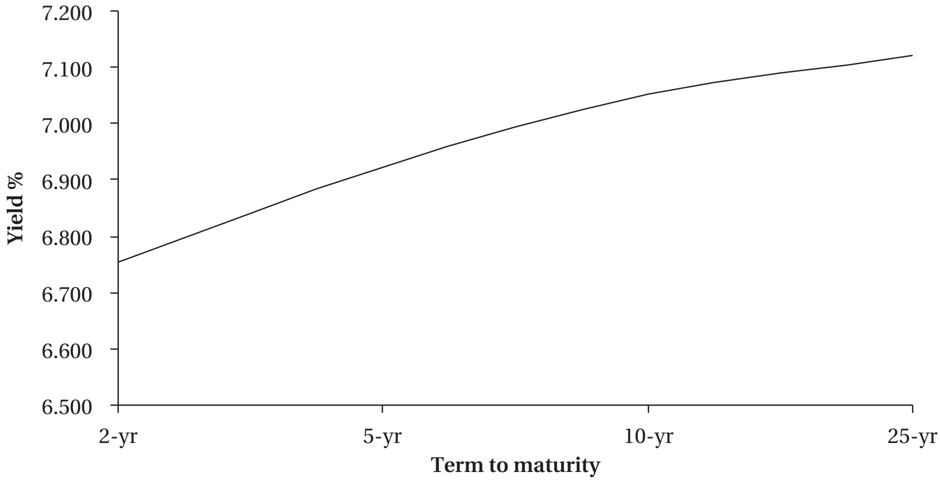


Figure 2.1: Gilt prices and yields for value 17 June 1997.
Source: Williams de Broe and Hambros Bank Limited; author's notes.

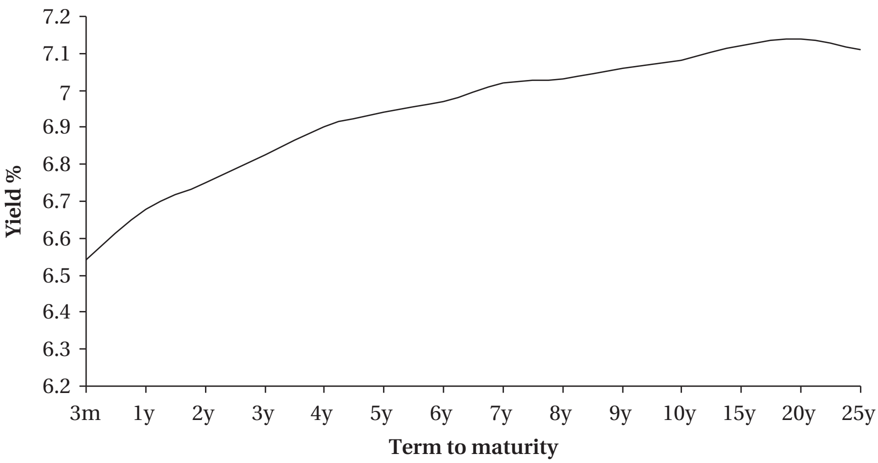


Figure 2.2: Benchmark gilt yield curve, 16 June 1997.
Source: Hambros Bank Limited; author's notes.

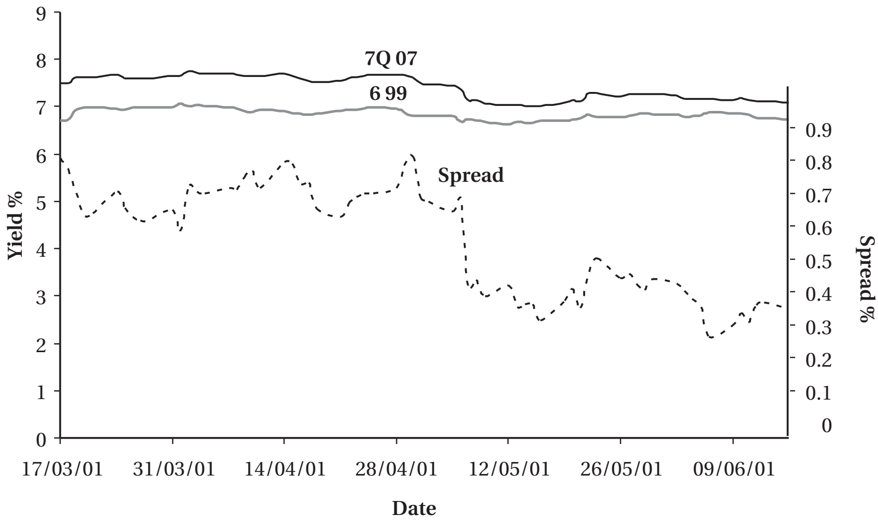


Figure 2.3: 6% Treasury 1999 and 7¼% 2007 three months' yield spread history as at 16 June 1997. Note: yield values are shown on the left axis, spread values on the right – which uses a larger scale for clarity.

(1230.04/166.42) or £7.3912 million of this bond, given by a simple ratio of the two basis point values. We expect to sell a greater nominal amount of the shorter-dated gilt because its risk exposure is lower. This trade generates cash because the short sale proceeds exceed the long buy purchase funds, which are, respectively,

Buy £1m 7.25% 2007	−£1,102,500
Sell £7.39m 6% 1999	+£7,437,025

What are the possible outcomes of this trade? If there is a parallel shift in the yield curve, the trade neither gains nor loses. If the yield spread narrows by, say, 15 basis points, the trade will gain either from a drop in yield on the long side or a gain in yield on the short side, or a combination of both. Conversely, a widening of the spread will result in a loss. Any narrowing spread is positive for the trade, while any widening is harmful.

The trade would be put on the same ratio if the amounts were higher, which is *scaling* the trade. So, for example, if the investor had bought £100 million of the 7.25% 2007, he would need to sell short £739 million of the two-year bonds. However, the risk exposure is greater by the same amount, so that in this case the trade would generate 100 times the risk. As can be imagined, there is a greater potential reward but at the same time a greater amount of stress in managing the position.

Using BPVs to risk-weight a relative value trade is common but suffers from any traditional duration-based measure because of the assumptions used in the analysis. Note that when using this method the ratio of the nominal amount of the bonds must equate the reciprocal of the bonds' BPV ratio. So in this case the BPV ratio is (166.42/1230.04) or 0.1353, which has a reciprocal of 7.3912. This means that the nominal values of the two bonds must always be in the ratio of 7.39:1. This weighting is not static, however; we know that duration measures are a static (snapshot) estimation of dynamic properties such as yield and term to maturity. Therefore, for anything but very short-term trades the relative values may need to be adjusted as the BPVs alter over time, so-called *dynamic adjustment* of the portfolio.

Another method to weight trades is by duration-weighting, which involves weighting in terms of market values. This compares to the BPV approach which provides a weighting ratio in terms of nominal values. In practice the duration approach does not produce any more accurate risk weighting.

A key element of any relative value trade is the financing cost of each position. This is where the repo market in each bond becomes important. In the example just described, the financing requirement is: repo out the 7.25% 2007, for which £1.1 million of cash must be borrowed to finance the purchase; the trader pays the repo rate on this stock reverse repo the 6% 1999 bond, which must be borrowed in repo to cover the short sale; the trader earns the repo rate on this stock.

If the repo rate on both stocks is close to the general repo rate in the market, there will be a bid–offer spread to pay but the greater amount of funds lent out against the 6% 1999 bond will result in a net financing gain on the trade whatever happens to the yield spread. If the 7.25% 2007 gilt is special, because the stock is in excessive demand in the market (for whatever reason), the financing gain will be greater still. If the 6% 1999 is *special*, the trade will suffer a financing loss.

In this case, however, the cash sums involved for each bond make the financing rates academic, as the amount paid in interest on the 7.25% 2007 repo is far outweighed by the interest earned on cash lent out when undertaking reverse repo in the 6% 1999 bond. Therefore, this trade will not be impacted by repo rate bid–offer spreads or specific rates, unless the rate on the borrowed bond is excessively special.[6] The repo financing cash flows for the 6% 1999 and 7.25% 2007 are shown in Figures 2.4 and 2.5 respectively, the Bloomberg repo/reverse repo screen RRRA. They show that at the time of the trade the investor had anticipated a 14-day term for the position before reviewing it and/or unwinding it.

A detailed account of the issues involved in financing a spread trade is contained in Choudhry (2002).[7]

2.1.3 *Identifying yield spread trades*

Yield spread trades are a type of relative value position that a trader can construct when the objective is to gain from a change in the spread between two points on the yield curve. The decision on which sectors of the curve to target is an important one and is based on a number of factors. An investor may naturally target, say, the five- and 10-year areas of the yield curve to meet investment objectives and have a view on these maturities. Or a trader may draw conclusions from studying the historical spread between two sectors.

Yield spreads do not move in parallel however and there is not a perfect correlation between the changes of short-, medium- and long-term sectors of the curve. The money market yield curve can sometimes act independently of the bond curve. Table 2.1 shows the change in benchmark yields during 1996/1997. There is no set pattern in the change in

[6] In fact the 6% 1999 did experience very special rates at certain times, briefly reaching negative rates at the start of 1998. However, the author had unwound the position long before then!

[7] Choudhry, M., *The Repo Handbook* (Oxford, UK: Butterworth-Heinemann, 2002). And just so you know, this trade was profitable as the yield spread between the 6% 1999 and 7.25% 2007 did indeed narrow, prior to the entire curve inverting just over one month later.

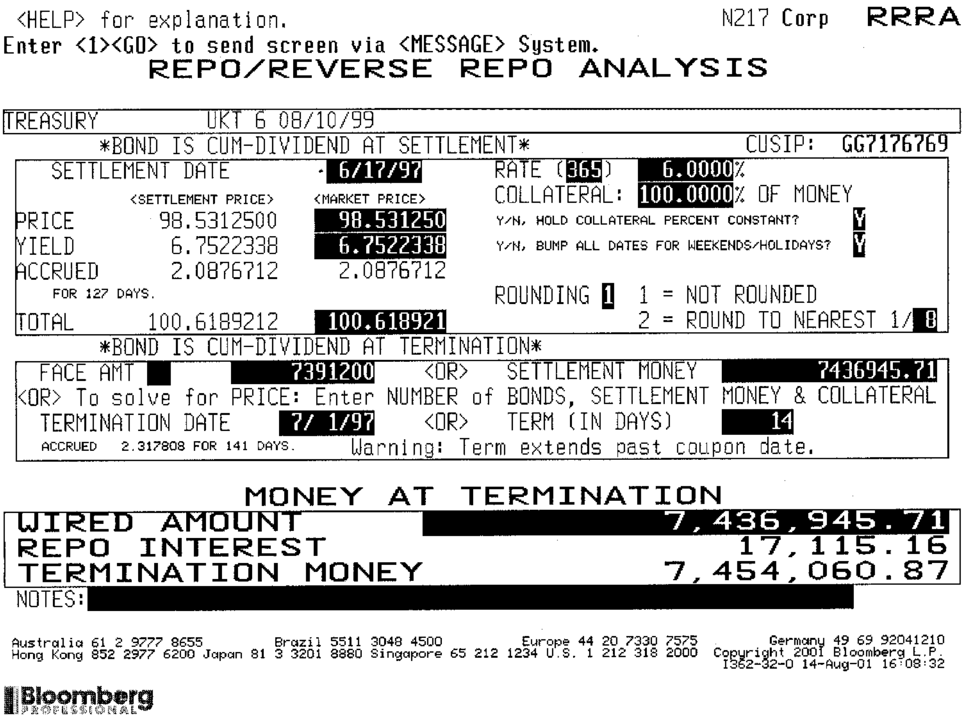


Figure 2.4: Bloomberg screen RRRA showing repo cash flows for 6% 1999, 17 June to 1 July 1997. ©Bloomberg L.P. Reproduced with permission.

both yield levels and spreads. It is apparent that one segment of the curve can flatten while another is steepening, or remains unchanged.

Another type of trade is where an investor has a view on one part of the curve relative to two other parts of the curve. This can be reflected in a number of ways, one of which is the butterfly trade, which is considered below.

2.2 Coupon spreads[8]

Coupon spreads are becoming less common in the gilt market because of the disappearance of high-coupon or other exotic gilts and the concentration on liquid benchmark issues. However, they are genuine spread trades. The US Treasury market presents greater opportunity for coupon spreads due to the larger number of similar-maturity issues. The basic principle behind the trade is a spread of two bonds that have similar maturity or similar duration but different coupons.

Table 2.2 shows the yields for a set of high-coupon and low(er)-coupon gilts for a specified date in May 1993 and the yields for the same gilts six months later. From the

[8] First presented by the author as an internal paper to the head of Treasury at ABN Amro Hoare Govett Sterling Bonds Limited in April 1995. Subsequently incorporated into this chapter.

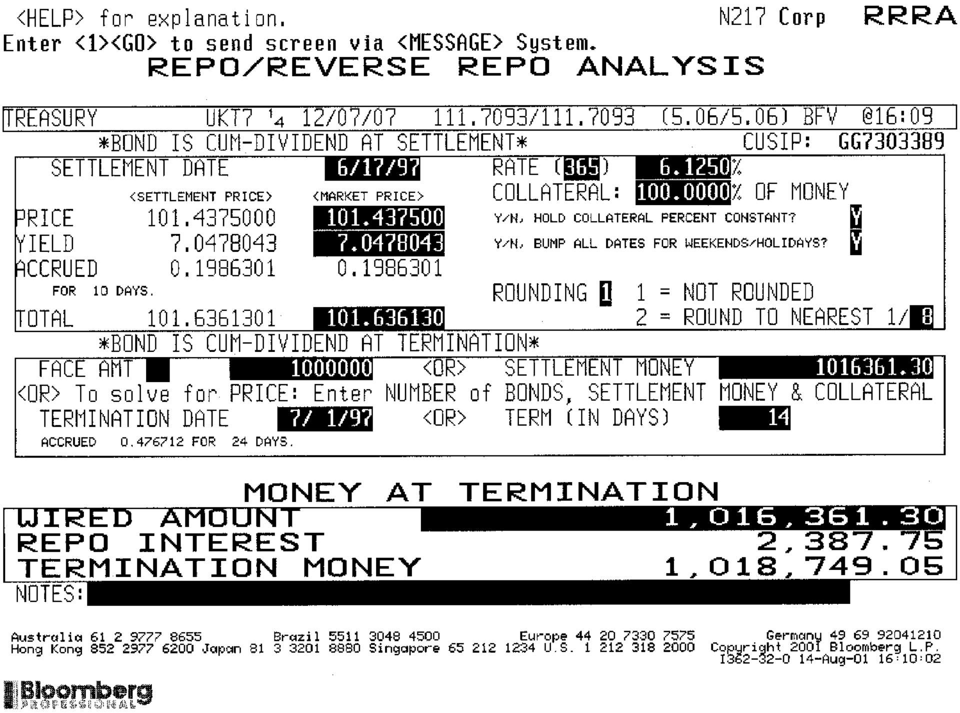


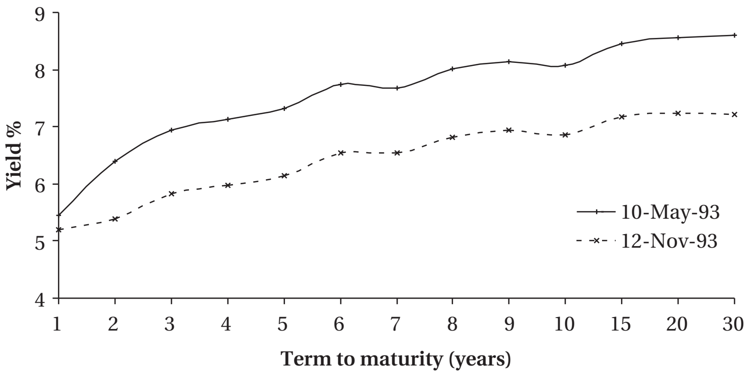
Figure 2.5: Bloomberg screen RRRA showing repo cash flows for 7.25% 2007, 17 June to 1 July 1997. ©Bloomberg L.P. Reproduced with permission.

	Changes in yield spread 3-month	**1-year**	**2-year**	**5-year**	**10-year**	**25-year**
10/11/1996	6.06	6.71	6.83	7.31	7.67	7.91
10/07/1997	6.42	6.96	7.057	7.156	7.025	6.921
	Change 0.36	0.25	0.227	−0.154	−0.645	−0.989
10/11/1997	7.15	7.3	7.09	6.8	6.69	6.47
	Change 0.73	0.34	0.033	−0.356	−0.335	−0.451

	Changes in yield spread 3m/1y	**1y/2y**	**2y/5y**	**5y/10y**	**5y/25y**	**10y/25y**
10/11/1996	−0.65	−0.12	−0.48	−0.36	−0.6	−0.24
10/07/1997	−0.54	−0.457	−0.099	0.131	0.235	0.104
	Change 0.11	−0.337	0.381	0.491	0.835	0.344
10/11/1997	−0.15	0.21	0.29	0.11	0.33	0.22
	Change 0.39	0.667	0.389	−0.021	0.095	0.116

Table 2.1: Yield levels and yield spreads during November 1996 to November 1997. Source: ABN Amro Hoare Govett Sterling Bonds Ltd, Hambros Bank Limited; Tullett & Tokyo; author's notes.

Stock	Term	10-May-93	12-Nov-93
Gilt	1	5.45	5.19
10Q 95	2	6.39	5.39
10 96	3	6.94	5.82
10H 97	4	7.13	5.97
9T 98 and 7Q 98	5	7.31	6.14
10Q 99	6	7.73	6.55
9 00	7	7.67	6.54
10 01	8	8.01	6.82
9T 02	9	8.13	6.95
8 03	10	8.07	6.85
9 08	15	8.45	7.18
9 12	20	8.55	7.23
8T 17	30	8.6	7.22



Gilt	Maturity	10/05/1993 Yield %	12/11/1993 Yield %	Yield change %
10Q 95	21-Jul-95	6.393	5.390	−1.003
14 96	10-Jan-96	6.608	5.576	−1.032
15Q 96	3-May-96	6.851	5.796	−1.055
13Q 96	15-May-96	6.847	5.769	−1.078
13Q 97	22-Jan-97	7.142	5.999	−1.143
10H 97	21-Feb-97	7.131	5.974	1.157
7 97	6-Aug-97	7.219	6.037	−1.182
8T 97	1-Sep-97	7.223	6.055	−1.168
15 97	27-Oct-97	7.294	6.113	−1.161
9T 98	19-Jan-98	7.315	6.102	−1.213
7Q 98	30-Mar-98	7.362	6.144	−1.218
6 99	10-Aug-99	7.724	6.536	−1.188
10Q 99	22-Nov-99	7.731	6.552	−1.179
8 03	10-Jun-03	8.075	6.854	−1.221
10 03	8-Sep-03	8.137	6.922	−1.215

Table 2.2: Yield changes on high- and low-coupon gilts from May 1993 to November 1993. Source: ABN Amro Hoare Govett Sterling Bonds Ltd; Bloomberg; author's notes.

yield curves we see that general yield levels decline by approximately 80–130 basis points. The last column in the table shows that, apart from the earliest pair of gilts (which do not have strictly comparable maturity dates), the performance of the lower-coupon gilt exceeded that of the higher-coupon gilt in every instance. Therefore, buying the spread of the low-coupon versus the high-coupon should, in theory, generate a trading gain in an environment of falling yields. One explanation of this is that the lower-coupon bonds are often the benchmark, which means the demand for them is higher. In addition, during a bull market, more bonds are considered to be 'high' coupon as overall yield levels decrease.

The exception noted in Table 2.2 is the outperformance of the 14% Treasury 1996 compared to the lower-coupon 10¼% 1995 stock. This is not necessarily conclusive, because the bonds are six months apart in maturity, which is a significant amount for short-dated stock. However, in an environment of low or falling interest rates, shorter-dated investors such as banks and insurance companies often prefer to hold very high-coupon bonds because of the high income levels they generate. This may explain the demand for the 14% 1996 stock[9] although the evidence at the time was only anecdotal.

2.3 Butterfly trades[10]

Butterfly trades are another method by which traders can reflect a view on changing yield levels without resorting to a naked punt on interest rates. They are another form of relative value trade; amongst portfolio managers they are viewed as a means of enhancing returns. In essence a butterfly trade is a short position in one bond against a long position of two bonds, one of shorter maturity and the other of longer maturity than the short-sold bond. Duration-weighting is used so that the net position is first-order risk-neutral, and nominal values are calculated such that the short sale and long purchase cash flows net to zero, or very closely to zero.

This section reviews some of the aspects of butterfly trades.

2.3.1 *Basic concepts*

A butterfly trade is par excellence a yield curve trade. If the average return on the combined long position is greater than the return on the short position (which is a cost) during the time the trade is maintained, the strategy will generate a profit. It reflects a view that the short-end of the curve will steepen relative to the 'middle' of the curve while the long-end will flatten. For this reason higher convexity stocks are usually preferred for the long positions, even if this entails a loss in yield. However, the trade is not 'risk-free', for the same reasons that a conventional two-bond yield spread is not. Although, in theory, a butterfly is risk-neutral with respect to parallel changes in the yield curve, changes in the shape of the curve can result in losses. For this reason the position must be managed dynamically and monitored for changes in risk relative to changes in the shape of the yield curve.

---

9 This stock also has a special place in the author's heart, although he was No. 2 on the desk when the Treasury head put on a very large position in it . . . !

10 Revised and updated version of a paper first presented internally to Adrian Howard (head of Cash-OBS desk, Treasury division) at Hambros Bank Limited in June 1997. Incorporated into this chapter.

In a butterfly trade the trader is long a short-dated and long-dated bond, and short a bond of a maturity that falls in between these two maturities. A portfolio manager with a constraint on running short positions may consider this trade as a switch out of a long position in the medium-dated bond and into duration-weighted amounts of the short-dated and long-dated bond. However, it is not strictly correct to view the combined long position to be an exact substitute for the short position – due to liquidity (and other reasons) the two positions will behave differently for given changes in the yield curve. In addition one must be careful to compare like for like, as the yield change in the short position must be analysed against yield changes in *two* bonds. This raises the issue of portfolio yield.

2.3.2 *Putting on the trade*

We begin by considering the calculation of the nominal amounts of the long positions, assuming a user-specified starting amount in the short position. In Table 2.3 we show three gilts as at 27 June 1997. The trade we wish to put on is a short position in the five-year bond, the 7% Treasury 2002, against long positions in the two-year bond, the 6% Treasury 1999 and the 10-year bond, the $7\frac{1}{4}$% Treasury 2007. Assuming £10 million nominal of the five-year bond, the nominal values of the long positions can be calculated using duration, modified duration or basis point values (the last two, unsurprisingly, will generate identical results). The more common approach is to use basis point values.

In a butterfly trade the net cash flow should be as close to zero as possible, and the trade must be basis point value-neutral. Using the following notation,

- P_1 the dirty price of the short position;
- P_2 the dirty price of the long position in the two-year bond;
- P_3 the dirty price of the long position in the 10-year bond;
- M_1 the nominal value of the short-position bond, with M_2 and M_3 the long-position bonds;
- BPV_1 the basis point value of the short-position bond;

if applying basis point values, the amounts required for each stock are given by

$$M_1P_1 = M_2P_2 + M_3P_3 \tag{2.1}$$

	2-year bond	5-year bond	10-year bond
Gilt	6% 1999	7% 2002	7.25% 2007
Maturity date	10 Aug 1999	07 Jun 2002	07 Dec 2007
Price	98-08	99-27	101-06
Accrued interest	2.30137	0.44110	0.45685
Dirty price	100.551	100.285	101.644
GRY %	6.913	7.034	7.085
Duration	1.969	4.243	7.489
Modified duration	1.904	4.099	7.233
Basis Point Value	0.01914	0.0411	0.07352
Convexity	0.047	0.204	0.676

Table 2.3: Bond values for butterfly strategy. Source: author's notes.

while the risk-neutral calculation is given by

$$M_1 BPV_1 = M_2 BPV_2 + M_3 BPV_3. \tag{2.2}$$

The value of M_1 is not unknown, as we have set it at £10 million. The equations can be rearranged to solve for the remaining two bonds, which are

$$\begin{aligned} M_2 &= \frac{P_1 BPV_3 - P_3 BPV_1}{P_2 BPV_3 - P_3 BPV_2} M_1 \\ M_3 &= \frac{P_2 BPV_1 - P_1 BPV_2}{P_2 BPV_3 - P_3 BPV_2} M_1. \end{aligned} \tag{2.3}$$

Using the dirty prices and BPVs from Table 2.3, we obtain the following values for the long positions. The position required is short £10 million 7% 2002 and long £5.347 million of the 6% 1999 and £4.576 million of the 7¼% 2007. With these values the trade results in a zero net cash flow and a first-order risk neutral interest-rate exposure. Identical results would be obtained using the modified duration values, and similar results using the duration measures. If using Macaulay duration the nominal values are calculated using

$$D_1 = \frac{MV_2 D_2 + MV_3 D_3}{MV_2 + MV_3} \tag{2.4}$$

where D and MV represent duration and market value for each respective stock.

2.3.3 *Yield gain*

We know that the gross redemption yield for a vanilla bond is that rate r where

$$P_d = \sum_{i=1}^{N} C_i e^{-rn}. \tag{2.5}$$

The right-hand side of equation (2.5) is simply the present value of the cash flow payments C to be made by the bond in its remaining lifetime. Equation (2.5) gives the continuously compounded yields to maturity; in practice users define a yield with compounding interval m, that is

$$r = (e^{rmn} - 1)/m. \tag{2.6}$$

Treasuries and gilts compound on a semi-annual basis.

In principle we may compute the yield on a portfolio of bonds exactly as for a single bond, using equation (2.5) to give the yield for a set of cash flows which are purchased today at their present value. In practice the market calculates portfolio yield as a weighted average of the individual yields on each of the bonds in the portfolio. This is described, for example, in Fabozzi (1993),[11] and this description points out the weakness of this method. An alternative approach is to weight individual yields using bonds' basis point values, which we illustrate here in the context of the earlier butterfly trade. In this trade we have

[11] Fabozzi, F., *Bond Portfolio Management*, Chapters 10–14 (FJF Associates, 1996).

- short £10 million 7% 2002;
- long £5.347 million 6% 1999 and £4.576 million 7¼% 2007.

Using the semi-annual adjusted form of equation (2.5) the true yield of the long position is 7.033%. To calculate the portfolio yield of the long position using market value weighting, we may use

$$r_{port} = \left(\frac{MV_2}{MV_{port}}\right) r_2 + \left(\frac{MV_3}{MV_{port}}\right) r_3 \tag{2.7}$$

which results in a portfolio yield for the long position of 6.993%. If we weight the yield with basis point values we use

$$r_{port} = \frac{BPV_2 M_2 r_2 + BPV_3 M_3 r_3}{BPV_2 M_2 + BPV_3 M_3}. \tag{2.8}$$

Substituting the values from Table 2.3 we obtain

$$\begin{aligned} r_{port} &= \frac{(1914)(5.347)(6.913) + (7352)(4.576)(7.085)}{(1914)(5.347) + (7352)(4.576)} \\ &= 7.045\%. \end{aligned}$$

We see that using basis point values produces a seemingly more accurate weighted yield, closer to the true yield computed using the expression above. In addition, using this measure a portfolio manager switching into the long butterfly position from a position in the 7% 2002 would pick up a yield gain of 1.2 basis points, compared to the 4 basis points that an analyst would conclude had been lost using the first yield measure.[12]

The butterfly trade therefore produces a yield gain in addition to the capital gain expected if the yield curve changes in the anticipated way.

2.3.4 *Convexity gain*

In addition to yield pick-up, the butterfly trade provides, in theory, a convexity gain which will outperform the short position irrespective of which direction interest rates move in, provided we have a parallel shift. This is illustrated in Table 2.4. This shows the changes in value of the 7% 2002 as interest rates rise and fall, together with the change in value of the combined portfolio.

We observe from Table 2.4 that whatever the change in interest rates, up to a point, the portfolio value will be higher than the value of the short position, although the effect is

[12] The actual income gained on the spread will depend on the funding costs for all three bonds, a function of the specific repo rates available for each bond. Shortly after the time of writing, the 6% Treasury 1999 went special, so the funding gain on a long position in this stock would have been excessive. However, buying the stock outright would have necessitated paying a yield premium, as demand for it increased as a result of it going special. In the event the premium was deemed high, an alternative stock was nominated, the 10¼% Conversion 1999, a bond with near-identical modified duration value.

Yield change (bps)	7% 2002 value (£)	Portfolio value* (£)	Difference (£)	BPV 7% 2002 (5-year)	BPV 6% 1999 (2-year)	BPV 7.25% 2007 (10-year)
+250	9,062,370	9,057,175	−5,195	0.0363	0.0180	0.0584
+200	9,246,170	9,243,200	−2,970	0.0372	0.0182	0.0611
+150	9,434,560	9,435,200	640	0.0381	0.0184	0.0640
+100	9,627,650	9,629,530	1,880	0.0391	0.0187	0.0670
+50	9,825,600	9,828,540	2,940	0.0401	0.0189	0.0702
0	10,028,500	10,028,500	0	0.0411	0.0191	0.0735
+50	10,236,560	10,251,300	14,740	0.0421	0.0194	0.0770
−100	10,450,000	10,483,800	33,800	0.0432	0.0196	0.0808
−150	10,668,600	10,725,700	57,100	0.0443	0.0199	0.0847
−200	10,893,000	10,977,300	84,300	0.0454	0.0201	0.0888
−250	11,123,000	11,240,435	117,435	0.0466	0.0204	0.0931

* Combined value of long positions in 6% 1999 and 7.25% 2007. Values rounded. Yield change is parallel shift.

Table 2.4: Changes in bond values with changes in yield levels.

progressively reduced as yields rise. The butterfly will always gain if yields fall, and protects against downside risk if yields rise to a certain extent. This is the effect of convexity; when interest rates rise, the portfolio value declines by less than the short position value, and when rates fall, the portfolio value increases by more. Essentially, the combined long position exhibits greater convexity than the short position. The effect is greater if yields fall, while there is an element of downside protection as yields rise, up to the +150 basis point parallel shift.

Portfolio managers may seek greater convexity whether or not there is a yield pick-up available from a switch. However, the convexity effect is only material for large changes in yield, and so if there was not a corresponding yield gain from the switch, the trade may not perform positively. As we noted, this depends partly on the funding position for each stock. The price/yield profile for each stock is shown in Figure 2.6.

Essentially, by putting on a butterfly as opposed to a two-bond spread or a straight directional play, the trader limits the downside risk if interest rates fall, while preserving the upside gain if yields fall.

2.3.5 *Analysis using Bloomberg screen 'BBA'*

To conclude the discussion of butterfly trade strategy, we describe the analysis using the 'BBA' screen on Bloomberg. The trade is illustrated in Figure 2.7.

Using this approach, the nominal values of the two long positions are calculated using BPV ratios only. This is shown under the column 'Risk Weight', and we note that the difference is zero. However, the nominal value required for the two-year bond is much greater, at £10.76 million, and for the 10-year bond much lower at £2.8 million. This results in a cash outflow of £3.632 million. The profit profile is, in theory, much improved; at the bottom of the screen we observe the results of a 100 basis point parallel shift in either direction, which is a profit. Positive results were also seen for 200 and 300 basis point parallel shifts in either direction. This screen incorporates the effect of a (uniform) funding

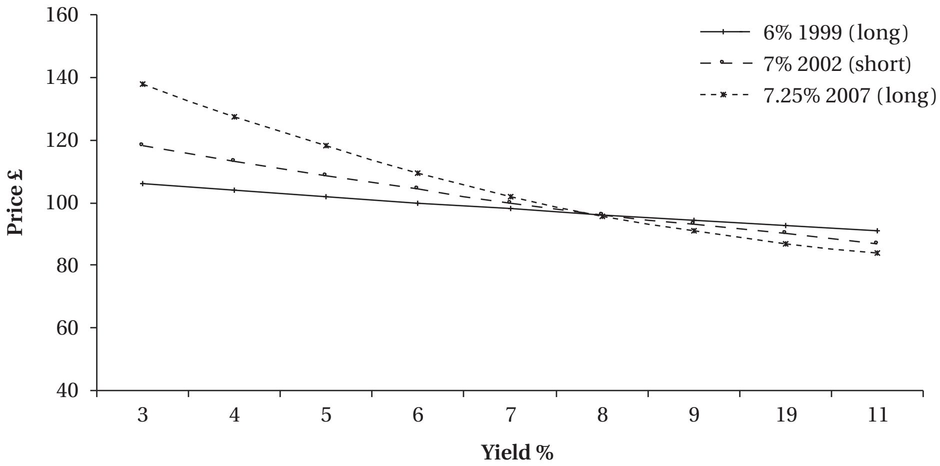


Figure 2.6: Illustration of convexity for each stock in butterfly trade, 27 June 1997.

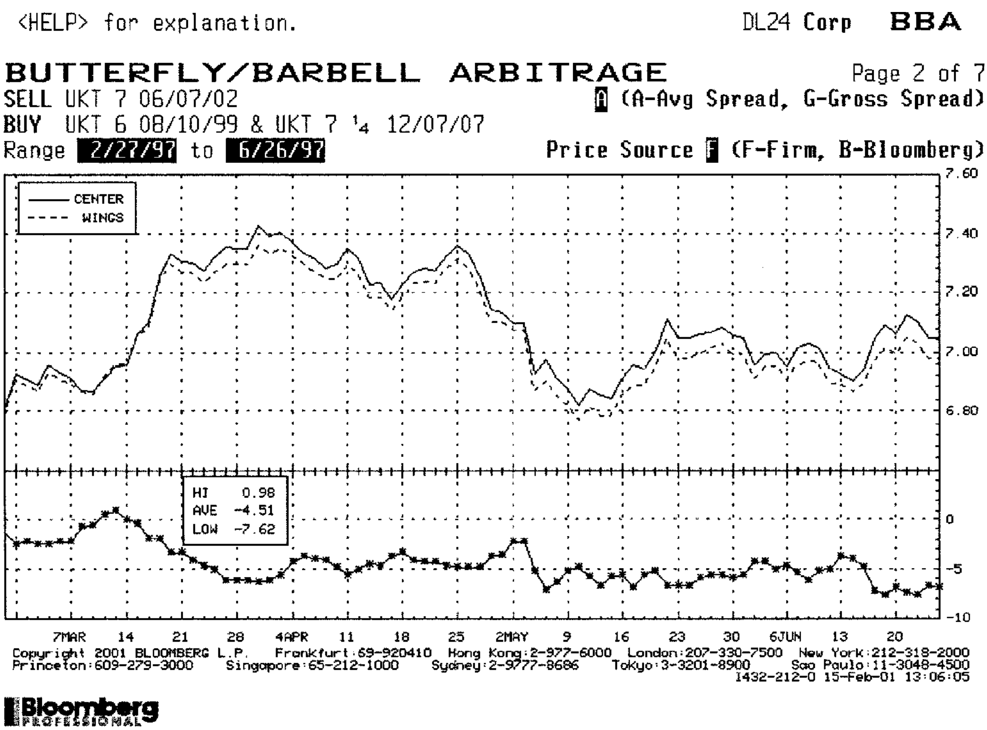


Figure 2.7: Butterfly trade analysis on 27 June 1997, on screen BBA. ©Bloomberg L.P. Reproduced with permission.

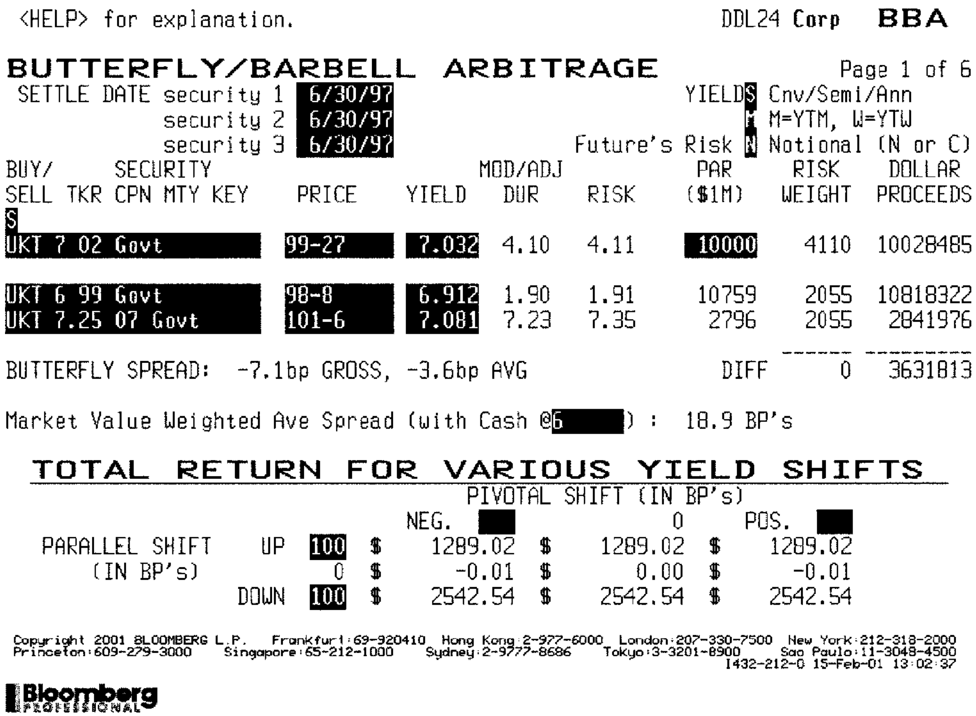


Figure 2.8: Butterfly trade spread history. ©Bloomberg L.P. Reproduced with permission.

rate, input on this occasion as 6.00%.[13] Note that the screen allows the user to see the results of a pivotal shift; however, in this example a 0 basis point pivotal shift is selected.

This trade therefore created a profit whatever direction interest rates moved in, assuming a parallel shift.

The spread history for the position up to the day before the trade is shown in Figure 2.8, a reproduction of the graph on Bloomberg screen BBA.

[13] In reality the repo rate will be slightly different for each stock, and there will be a bid–offer spread to pay, but as long as none of the stocks are special the calculations should be reasonably close.

3 The Dynamics of Asset Prices

The modelling of the yield curve is a function of the movement in the price of the underlying asset, which in this case is the movement in interest rates. Both option valuation models and interest-rate models describe an environment where the price of an option (or the modelling of the yield curve) is related to the behaviour process of the variables that drive asset prices. This process is described as a *stochastic* process, and pricing models describe the stochastic dynamics of asset price changes, whether this is change in share prices, interest rates, foreign exchange rates or bond prices. To understand the mechanics of option pricing therefore, we must familiarise ourselves with the behaviour of functions of *stochastic variables*. The concept of a stochastic process is a vital concept in finance theory. It describes random phenomena that evolve over time, and these include asset prices. For this reason an alternative title for this chapter could be *An Introduction to Stochastic Processes*.

This is a book on bonds after all, not mathematics, and it is outside the scope of this book comprehensively to derive and prove the main components of dynamic asset pricing theory. There are a number of excellent textbooks that the reader is encouraged to read which provide the necessary detail, in particular Ingersoll (1987), Baxter and Rennie (1996), Neftci (1996) and James and Webber (2000). Another recommended text that deals with probability models in general, as well as their application in derivatives pricing, is Ross (2000). In this chapter we review the basic principles of the dynamics of asset prices, which are required for an understanding of interest-rate modelling. The main principles are then considered again in the context of yield curve modelling, in the following chapters.

3.1 The behaviour of asset prices

The first property that asset prices, which can be taken to include interest rates, are assumed to follow is that they are part of a *continuous* process. This means that the value of any asset can and does change at any time and from one point in time to another, and can assume any fraction of a unit of measurement. It is also assumed to pass through every value as it changes, so for example if the price of a bond moves from 92.00 to 94.00 it must also have passed through every point in between. This feature means that the asset price does not exhibit *jumps*, which in fact is not the case in many markets, where price processes do exhibit jump behaviour. For now however we may assume that the price process is continuous.

3.1.1 *Stochastic processes*

Models that seek to value options or describe a yield curve also describe the dynamics of asset price changes. The same process is said to apply to changes in share prices, bond prices, interest rates and exchange rates. The process by which prices and interest rates evolve over time is known as a *stochastic process*, and this is a fundamental concept in finance theory.[1] Essentially a stochastic process is a time series of random variables. Generally the random variables in a stochastic process are related in a non-random manner, and so therefore we can capture them in a *probability density function*. A good introduction is given in Neftci (1996), and following his approach we very briefly summarise the main features here.

Consider the function $y = f(x)$; given the value of x we can obtain the value of y. If we denote the set W as the state of the world, where $w \in W$, the function $f(x, w)$ has the property that given a value $w \in W$ it becomes a function of x only. If we say that x represents the passage of time, two functions $f(x, w_1)$ and $f(x, w_2)$ will be different because the second element w in each case is different. With x representing time, these two functions describe two different processes that are dependent on different states of the world W. The element w represents an underlying random process, and so therefore the function $f(x, w)$ is a *random function*. A random function is also called a *stochastic process*, one in which x represents time and $x \geq 0$. The random characteristic of the process refers to the entire process, and not any particular value in that process at any particular point in time.

Examples of functions include the *exponential* function denoted by $y = e^x$ and the *logarithmic* function $\log_e(y) = x$.

The price processes of shares and bonds, as well as interest rate processes, are stochastic processes. That is, they exhibit a random change over time. For the purposes of modelling, the change in asset prices is divided into two components. These are the *drift* of the process, which is a *deterministic* element,[2] also called the mean, and the random component known as the *noise*, also called the volatility of the process.

We introduce the drift component briefly as follows. For an asset such as an ordinary share, which is expected to rise over time (at least in line with assumed growth in inflation), the drift can be modelled as a geometric growth progression. If the price process had no 'noise', the change in price of the stock over the time period dt can be given by

$$\frac{\mathrm{d}S_t}{\mathrm{d}t} = \mu S_t \tag{3.1}$$

where the term μ describes the growth rate. Expression (3.1) can be rewritten in the form

$$\mathrm{d}S_t = \mu S_t \mathrm{d}t \tag{3.2}$$

which can also be written in integral form. For interest rates, the movement process can be described in similar fashion, although as we shall see interest rate modelling often takes into account the tendency for rates to return to a mean level or range of levels, a process known as *mean reversion*. Without providing the derivation here, the equivalent expression for interest rates takes the form

$$\mathrm{d}r_t = \alpha(\mu - r_t)\mathrm{d}t \tag{3.3}$$

1 A formal definition of a stochastic process is given in Appendix 3.1.

2 There are two types of model: *deterministic*, which involves no randomness so the variables are determined exactly; and *stochastic*, which incorporates the random nature of the variables into the model.

where α is the mean reversion rate that determines the pace at which the interest rate reverts to its mean level. If the initial interest rate is less than the drift rate, the rate r will increase, while if the level is above the drift rate it will tend to decrease.

For the purposes of employing option pricing models the dynamic behaviour of asset prices is usually described as a function of what is known as a *Weiner process*, which is also known as *Brownian motion*. The noise or volatility component is described by an *adapted* Brownian or Weiner process, and involves introducing a random increment to the standard random process. This is described next.

3.1.2 *Weiner process or Brownian motion*

The stochastic process we have briefly discussed above is known as Brownian motion or a Weiner process. In fact a Weiner process is only a process that has a mean of 0 and a variance of 1, but it is common to see these terms used synonymously. Weiner processes are a very important part of continuous-time finance theory, and interested readers can obtain more detailed and technical data on the subject in Neftci (1996) and Duffie (1996)[3] among others. It is a well-researched subject.

One of the properties of a Weiner process is that the sample pathway is continuous, that is, there are no *discontinuous* changes. An example of a discontinuous process is the Poisson process. Both are illustrated in Figures 3.1 and 3.2 below.

In the examples illustrated, both processes have an expected change of 0 and a variance of 1 per unit of time. There are no discontinuities in the Weiner process, which is a plot of many very tiny random changes. This is reflected in the 'fuzzy' nature of the sample path. However the Poisson process has no fuzzy quality and appears to have a much smaller number of random changes. We can conclude that asset prices, and the dynamics of interest rates, are more akin to a Weiner process. This, therefore is how asset prices are modelled. From observation we know that, in reality asset prices and interest rates do exhibit discontinuities or *jumps*, however there are other advantages to assuming a Weiner process, and in practice because continuous-time stochastic processes can be captured as a combination of Brownian motion and a Poisson process, analysts and researchers use the former as the basis of financial valuation models.

The first step in asset pricing theory builds on the assumption that prices follow a Brownian motion. The properties of Brownian motion W state that it is continuous, and the value of W_t $(t > 0)$ is normally distributed under a probability measure P as a *random*

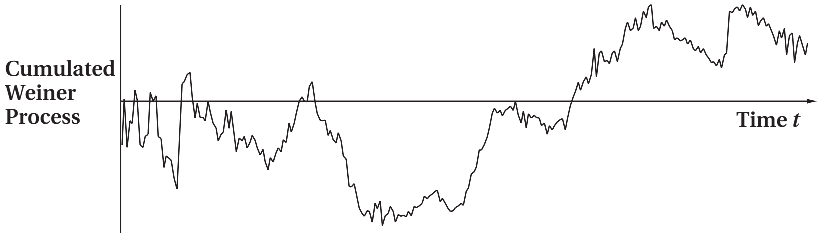


Figure 3.1: An example of a Weiner process.

[3] Duffie's text requires a very good grounding in continuous-time mathematics.

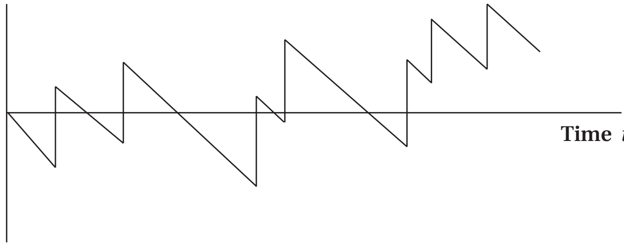


Figure 3.2: An example of a Poisson process.

variable with parameters $N(0, t)$. An incremental change in the asset value over time dt, which is a very small or *infinitesimal* change in the time, given by $W_{s+t} - W_s$, is also normally distributed with parameters $N(0, t)$ under P. Perhaps the most significant feature is that the change in value is independent of whatever the history of the price process has been up to time s. If a process follows these conditions it is Brownian motion. In fact asset prices do not generally have a mean of 0, because over time we expect them to rise. Therefore modelling asset prices incorporates a *drift* measure that better reflects asset price movement, so that an asset movement described by

$$S_t = W_t + \mu t \tag{3.4}$$

would be a Brownian motion with a drift given by the constant μ. A second parameter is then added, a *noise* factor, which scales the Brownian motion by another constant measure, the standard deviation σ. The process is then described by

$$S_t = \sigma W_t + \mu t \tag{3.5}$$

which can be used to *simulate* the price path taken by an asset, as long as we specify the two parameters. An excellent and readable account of this is given in Baxter and Rennie (1996, Chapter 3), who also state that under (3.5) there is a possibility of achieving negative values, which is not realistic for asset prices. However using the exponential of the process given by (3.5) is more accurate, and is given by (3.6):

$$S_t = exp(\sigma W_t + \mu t). \tag{3.6}$$

Brownian motion or the *Weiner process* is employed by virtually all option pricing models, and we introduce it here with respect to a change in the variable W over an interval of time t. If W represents a variable following a Weiner process and ΔW is a change in value over a period of time t, the relationship between ΔW and Δt is given by (3.7):

$$\Delta W = \varepsilon\sqrt{\Delta t} \tag{3.7}$$

where ε is a random sample from a normal distribution with a mean 0 and a standard deviation of 1. Over a short period of time the values of ΔW are independent and therefore also follow a normal distribution with a mean of 0 and a standard deviation of $\sqrt{\Delta t}$. Over a longer time period T made up of N periods of length Δt, the change in W over the period from time 0 to time T is given by (3.8):

$$W(T) - W(0) = \sum_{i=1}^{N} \varepsilon_i\sqrt{\Delta t}. \tag{3.8}$$

The successive values assumed by W are serially independent so from (3.8) we conclude that changes in the variable W from time 0 to time T follow a normal distribution with mean 0 and a standard deviation of $\sqrt{T}$. This describes the Weiner process, with a mean of zero or a zero drift rate and a variance of T. This is an important result because a zero drift rate implies that the change in the variable (for which now read asset price) in the future is equal to the current change. This means that there is an equal chance of an asset return ending up 10% or down 10% over a long period of time.

The next step in the analysis involves using stochastic calculus. Without going into this field here, we summarise from Baxter and Rennie (1996) and state that a stochastic process X will incorporate a *Newtonian* term that is based on dt and a Brownian term based on the infinitesimal increment of W that is denoted by $\mathrm{d}W_t$. The Brownian term has a 'noise' factor of σ_t. The infinitesimal change of X at X_t is given by the differential equation

$$\mathrm{d}X_t = \sigma_t \mathrm{d}W_t + \mu_t \mathrm{d}t \tag{3.9}$$

where σ_t is the *volatility* of the process X at time t and μ_t is the drift of X at time t. For interest rates that are modelled on the basis of mean reversion, the process is given by

$$\mathrm{d}r_t = \sigma_t \mathrm{d}W_t + \alpha(u_t - r_t)\mathrm{d}t \tag{3.10}$$

where the mean reverting element is as before. Without providing the supporting mathematics, which we have not covered here, the process described by (3.10) is called an Ornstein–Uhlenbeck process, and has been assumed by a number of interest rate models.

One other important point to introduce here is that a random process described by (3.10) operates in a continuous environment. In continuous-time mathematics the *integral* is the tool that is used to denote the sum of an infinite number of objects, that is where the number of objects is *uncountable*. A formal definition of the integral is outside the scope of this book, but accessible accounts can be found in the texts referred to previously. A basic introduction is given in Appendix 3.4. However the continuous stochastic process X described by (3.9) can be written as an integral equation in the form

$$X_t = X_0 + \int_0^t \sigma_s \mathrm{d}W_s + \int_0^t \mu_s \mathrm{d}s \tag{3.11}$$

where σ and μ are processes as before. The volatility and drift terms can be dependent on the time t but can also be dependent on X or W up to the point t. This is a complex technical subject and readers are encouraged to review the main elements in the referred texts.

3.1.3 *The martingale property*

Continuous time asset pricing is an important part of finance theory and involves some quite advanced mathematics. An excellent introduction to this subject is given in Baxter and Rennie (1996) and Neftci (1996). A more technical account is given in Williams (1991). It is outside the scope of this book to derive, prove and detail the main elements. However we wish to summarise the essential property, and begin by saying that in continuous time, asset prices can take on an unlimited number of values. Stochastic differential equations are used to capture the dynamics of asset prices in a generalised form. So for example, as we saw in the previous section an incremental change in the price of an asset S at time t could be given by

$$\mathrm{d}S = \mu S \mathrm{d}t + \sigma S \mathrm{d}W(t) \tag{3.12}$$

where

$\mathrm{d}S$	is an infinitesimal change in the price of asset S
$\mu S\mathrm{d}t$	is the predicted movement during the infinitesimal time interval $\mathrm{d}t$
$\sigma S\mathrm{d}W(t)$	is an unpredictable random shock.

Martingale theory is a branch of mathematics that classifies the *trend* in an observed time series set of data. A stochastic process is said to behave like a martingale if there are no observable trends in its pattern. The martingale property is often used in conjunction with a Weiner process to describe asset price dynamics. The notion of the martingale property is that the best approximation of a set of integrable random variables M at the end of a time period t is M_0, which essentially states that the most accurate way to predict a future asset price is to use the price of the asset now. That is, using the price today is the same as using all available historical information, as only the newest information regarding the asset is relevant.

We do not describe or prove this property here but the martingale property is used to derive (3.13), the price of an asset at time t:

$$P_t = exp\left(\sigma W_t - \tfrac{1}{2}\sigma^2 t\right). \tag{3.13}$$

A martingale is an important type of stochastic process and the concept of a martingale is fundamental to asset pricing theory. A process that is a martingale is one in which the expected future value, based on what is known up to now, is the same as today's value. That is a martingale is a process in which the *conditional* expected future value, given current information, is equal to the current value. The martingale representation theorem states that given a Weiner process, and the fact that the path of the Weiner process up to that point is known, then any martingale is equal to a constant plus a stochastic integral, with respect to the Weiner process. This can be written as

$$E_t[S_T] = S_t \text{ for } t \leq T. \tag{3.14}$$

Therefore a stochastic process that is a martingale has no observable *trend*. The price process described by (3.9) is not a martingale unless the drift component μ is equal to zero, otherwise a trend will be observed. A process that is observed to trend upwards is known as a *submartingale*, while a process that on average declines over time is known as a *supermartingale*.

What is the significance of this? Here we take it as given that because price processes can be described as *equivalent martingale measures* (which we do not go into here) they enable the practitioner to construct a risk-free hedge of a market instrument. By enabling a no-arbitrage portfolio to be described, a mathematical model can be set up and solved, including risk-free valuation models.

The background and mathematics to martingales can be found in Harrison and Kreps (1979) and Harrison and Pliska (1981) as well as Baxter and Rennie (1996). For a description of how, given that price processes are martingales, we are able to price derivative instruments see James and Webber (2000, Chapter 4).

3.1.4 *Generalised Weiner process*

The standard Weiner process is a close approximation of the behaviour of asset prices but does not account for some specific aspects of market behaviour. In the first instance the prices of financial assets do not start at zero, and their price increments have positive mean. The variance of asset price moves is also not always unity. Therefore the standard Weiner process is replaced by the generalised Weiner process, which describes a variable that may start at something other than zero, and also has incremental changes that have a mean other than zero as well as variances that are not unity. The mean and variance are still constant in a generalised process, which is the same as the standard process, and a different description must be used to describe processes that have variances that differ over time; these are known as stochastic integrals.

We now denote the variable as X and for this variable a generalised Weiner process is given by (3.15):

$$\mathrm{d}X = a\mathrm{d}t + b\mathrm{d}W \tag{3.15}$$

where a and b are constants. This expression describes the dynamic process of the variable X as a function of time and $\mathrm{d}W$. The first term $a\mathrm{d}t$ is known as the deterministic term and states that the expected drift rate of X over time is a per unit of time; the second term $b\mathrm{d}W$ is the stochastic element and describes the variability of the move in X over time, and is quantified by b multiplied by the Weiner process. When the stochastic element is zero, $\mathrm{d}X = a\mathrm{d}t$, or put another way

$$\mathrm{d}X/\mathrm{d}t = a.$$

From this we state that at time 0, $X = X_0 + at$. This enables us to describe the price of an asset, given its initial price, over a period of time. That is, the value of X at any time is given by its initial value at time 0, which is X_0, together with its drift multiplied by the length of the time period. We can restate (3.15) to apply over a long time period Δt, shown as (3.16):

$$\Delta X = a\Delta t + b\varepsilon\sqrt{\mathrm{d}t}. \tag{3.16}$$

As with the standard Weiner process ΔX has a normal distribution with mean $a\Delta t$ and standard deviation $b\sqrt{\Delta t}$.

The generalised Weiner process is more flexible than the standard one but is still not completely accurate as a model of the behaviour of asset prices. It has normally distributed

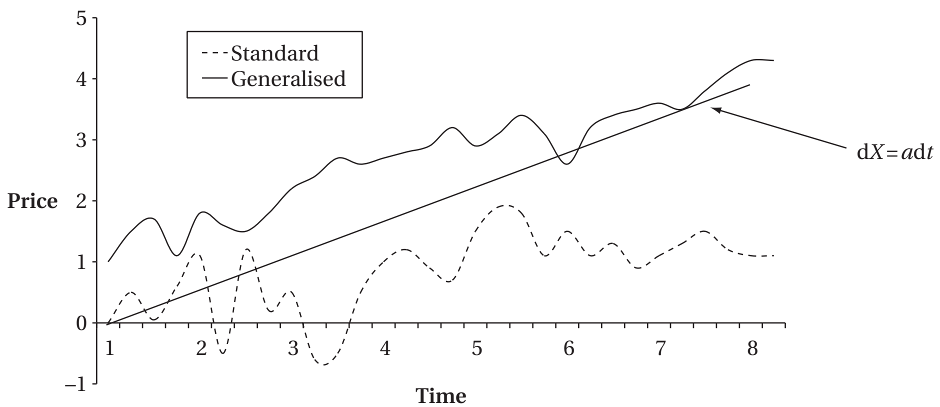


Figure 3.3: Standard and generalised Weiner processes.

values, which means that there is a probability of observing negative prices. For assets such as equities, this is clearly unrealistic. In addition the increments of a Weiner process are additive whereas the increments of asset prices are more realistically multiplicative. In fact as the increments of a Weiner process have constant expectation, this implies that the percentage incremental change in asset prices, or the percentage rate of return on the stock, would be declining as the stock price rises. This is also not realistic. For this reason a geometric process or geometric Brownian motion has been introduced,[4] which is developed by an exponential transformation of the generalised process. From (3.16), a one-dimensional process is a geometric Brownian motion if it has the form e^X, where X is a one-dimensional generalised Brownian motion with a deterministic initial value of $X(0)$.

Another type of stochastic process is an Itô process. This a generalised Weiner process where the parameters a and b are functions of the value of the variable X and time t. An Itô process for X can be written as (3.17):

$$dX = a(X, t)dt + b(X, t)dW. \tag{3.17}$$

The expected drift rate and variance of an Itô process are liable to change over time; indeed the dependence of the expected drift rate and variance on X and t is the main difference between it and a generalised Weiner process. The derivation of Itô's formula is given in Appendix 3.3.

3.1.5 *A model of the dynamics of asset prices*

The above discussion is used to derive a model of the behaviour of asset prices sometimes referred to as *geometric Brownian motion*. The dynamics of the asset price X are represented by the Itô process shown in (3.18), where there is a drift rate of a and a variance rate of b^2X^2,

$$dX = aXdt + bXdW \tag{3.18}$$

so that

$$\frac{dX}{X} = a + bdW.$$

The uncertainty element is described by the Weiner process element, with

$$dW = \varepsilon\sqrt{dt}$$

where ε is the error term, a random sample from the standardised normal distribution, so that $\varepsilon \sim N(0, 1)$. From this, and over a longer period of time Δt we can write

$$\frac{\Delta X}{X} = a\Delta t + b\varepsilon\sqrt{\Delta t}.$$

Over this longer period of time, for application in a discrete-time environment, if we assume that volatility is zero, we have

$$\Delta X = a\Delta t + b\varepsilon\sqrt{\Delta t} \tag{3.19}$$

4 See for instance, Nielson (1999).

and

$$dX = a dt \quad \text{and} \quad \frac{dX}{dt} = a.$$

Return is given by $X = X_0 e^{at}$.

The discrete time version of the asset price model states that the proportional return on the asset price X over a short time period is given by an expected return of $a\Delta t$ and a stochastic return of $b\varepsilon\Delta dt$. Therefore the returns of asset price changes $\Delta X/X$ are normally distributed with a mean of $a\Delta t$ and a standard deviation of $b\sqrt{\Delta t}$. This is the distribution of asset price returns and is given by (3.20):

$$\frac{\Delta X}{X} \sim N(a\Delta t,\ b\sqrt{\Delta t}) \tag{3.20}$$

Example 3.1

A conventional bond has an expected return of 5.875% and a standard deviation of 12.50% per annum. The initial price of the bond is 100. From (3.20) the dynamics of the bond price are given by:

$$dP/P = 0.05875dt + 0.125dW$$

and for a time period Δt by $dP/P = 0.05875\Delta t + 0.125\varepsilon\sqrt{\Delta t}$.

If the short time interval Δt is four weeks or 0.07692 years, assuming $\varepsilon = 1$, then the increase in price is given by:

$$\begin{aligned}\Delta P &= 100(0.05875(0.07692) + 0.125\varepsilon\sqrt{0.07692})\\ &= 100(0.00451905 + 0.0346681\varepsilon).\end{aligned}$$

So the price increase is described as a random sample from a normal distribution with a mean of 0.452 and a volatility of 3.467. Over a time interval of four weeks $\Delta P/P$ is normal with:

$$\Delta P/P \sim N(0.00452, 0.001202).$$

3.1.6 *The distribution of the risk-free interest rate*

The continuously compounded rate of return is an important component of option pricing theory. If r is the continuously compounded rate of return, we can use the lognormal property to determine the distribution that this follows. At a future date T the asset price S may be written as (3.21):

$$S_T = S_t e^{r(T-t)} \tag{3.21}$$

and $r = \dfrac{1}{T-t} ln\left(\dfrac{S_T}{S_t}\right)$.

Using the lognormal property we can describe the distribution of the risk-free rate as:

$$r \sim N\left((\mu - \tfrac{1}{2}\sigma^2),\ \frac{\sigma}{\sqrt{T-t}}\right). \tag{3.22}$$

3.2 Stochastic calculus models: Brownian motion and Itô calculus

We noted at the start of the chapter that the price of an option is a function of the price of the underlying stock and its behaviour over the life of the option. Therefore this option price is determined by the variables that describe the process followed by the asset price over a continuous period of time. The behaviour of asset prices follows a stochastic process, and so option pricing models must capture the behaviour of stochastic variables behind the movement of asset prices. To accurately describe financial market processes a financial model will depend on more than one variable. Generally a model is constructed where a function is itself a function of more than one variable. Itô's lemma, the principal instrument in continuous time finance theory, is used to differentiate such functions. This was developed by a mathematician, K. Itô, in 1951. Here we simply state the theorem, as a proof and derivation are outside the scope of the book. Interested readers may wish to consult Briys *et al.* (1998) and Hull (1997) for a background on Itô's lemma; we also recommend Neftci (1996). Basic background on Itô's lemma is given in Appendices 3.2 and 3.3.

3.2.1 *Brownian motion*

Brownian motion is very similar to a Weiner process, which is why it is common to see the terms used interchangeably. Note that the properties of a Weiner process require that it be a martingale, while no such constraint is required for a Brownian process. A mathematical property known as the *Lévy theorem* allows us to consider any Weiner process W_t with respect to an information set F_t as a Brownian motion Z_t with respect to the same information set.

We can view Brownian motion as a continuous time *random walk*, visualised as a walk along a line, beginning at $X_0 = 0$ and moving at each incremental time interval $\mathrm{d}t$ either up or down by an amount $\sqrt{\mathrm{d}t}$. If we denote the position of the walk as X_n after the nth move, the position would be

$$X_n = X_{n-1} \pm \sqrt{\mathrm{d}t}, \qquad n = 1, 2, 3 \ldots \tag{3.23}$$

where the $+$ and $-$ signs occur with an equal probability of 1/2. This is a simple random walk. We can transform this into a continuous path by applying linear interpolation between each move point, so that

$$\bar{X}_t = X_n + (t - n\mathrm{d}t) \cdot (X_{n+1} - X_n), \qquad n\mathrm{d}t \leq t \leq (n+1)\mathrm{d}t. \tag{3.24}$$

It can be shown (but not here) that the path described in (3.24) has a number of properties, including that the incremental change in value each time it moves is independent of the behaviour leading up to the move, and that the mean value is 0 and variance is finite. The mean and variance of the set of moves is independent of $\mathrm{d}t$.

What is the importance of this? Essentially this: the probability distribution of the motion can be shown, as $\mathrm{d}t$ approaches 0, to be normal or *Gaussian*.

3.2.2 *Stochastic calculus*

Itô's theorem provides an analytical formula that simplifies the treatment of stochastic differential equations, which is why it is so valuable. It is an important rule in the application of stochastic calculus to the pricing of financial instruments. Here we briefly describe the power of the theorem.

The standard stochastic differential equation for the process of an asset price S_t is given in the form

$$\mathrm{d}S_t = a(S_t,\ t)\mathrm{d}t + b(S_t,\ t)\mathrm{d}W_t \tag{3.25}$$

where $a(S_t,\ t)$ is the drift coefficient and $b(S_t,\ t)$ is the volatility or *diffusion* coefficient. The Weiner process is denoted $\mathrm{d}W_t$ and refers to the unpredictable events that occur at time intervals $\mathrm{d}t$. This is sometimes denoted $\mathrm{d}Z$ or $\mathrm{d}z$.

Consider a function $f(S_t,\ t)$ dependent on two variables S and t, where S follows a random process and varies with t. If S_t is a continuous-time process that follows a Weiner process W_t, then it directly influences the function $f(\ )$ through the variable t in $f(S_t,\ t)$. Over time we observe new information about W_t as well as the movement in S over each time increment, given by $\mathrm{d}S_t$. The sum of both these effects represents the *stochastic differential* and is given by the stochastic equivalent of the chain rule known as *Itô's lemma*. So for example, if the price of a stock is 30 and an incremental time period later is 30½, the differential is ½.

If we apply a Taylor expansion in two variables to the function $f(S_t,\ t)$ we obtain

$$\mathrm{d}f_t = \frac{\partial f}{\partial S_t}\mathrm{d}S_t + \frac{\partial f}{\partial t}\mathrm{d}t + \frac{1}{2}\frac{\partial^2 f}{\partial S_t^2}b_t^2\mathrm{d}t. \tag{3.26}$$

Remember that ∂t is the partial derivative while $\mathrm{d}t$ is the derivative.

If we substitute the stochastic differential equation (3.25) for S_t we obtain *Itô's lemma* of the form

$$\mathrm{d}f_t = \left(\frac{\partial f}{\partial S_t}a_t + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial S_t^2}b_t^2\right)\mathrm{d}t + \frac{\partial f}{\partial S_t}b_t\mathrm{d}W_t. \tag{3.27}$$

What we have done is taken the stochastic differential equation ('SDE') for S_t and transformed it so that we can determine the SDE for f_t. This is absolutely priceless, a valuable mechanism by which we can obtain an expression for pricing derivatives that are written on an underlying asset whose price can be determined using conventional analysis. In other words, using Itô's formula enables us to determine the SDE for the derivative, once we have set up the SDE for the underlying asset. This is the value of Itô's lemma.

The SDE for the underlying asset S_t is written in most textbooks in the following form:

$$\mathrm{d}S_t = \mu S_t\mathrm{d}t + \sigma S_t\mathrm{d}W_t \tag{3.28}$$

which has simply denoted the drift term $a(S_t,\ t)$ as μS_t and the diffusion term $b(S_t,\ t)$ as σS_t. In the same way Itô's lemma is usually seen in the form

$$\mathrm{d}F_t = \left[\frac{\partial F}{\partial S_t}\mu S_t + \frac{\partial F}{\partial t} + \frac{1}{2}\frac{\partial^2 F}{\partial S_t^2}\sigma^2 S_t^2\right]\mathrm{d}t + \frac{\partial F}{\partial S_t}\sigma S_t\mathrm{d}W_t \tag{3.29}$$

although the noise term is sometimes denoted $\mathrm{d}Z$. Further applications are illustrated in Example 3.2.

Example 3.2
(i) Lognormal distribution
A variable (such as an asset price) may be assumed to have a *lognormal distribution* if the natural logarithm of the variable is normally distributed. So if an asset price S follows a stochastic process described by

$$\mathrm{d}S = \mu S\mathrm{d}t + \sigma S\mathrm{d}W \tag{3.30}$$

how would we determine the expression for *ln S*? This can be achieved using Itô's lemma.

If we say that $F = ln\ S$, then the first derivative

$\frac{\mathrm{d}F}{\mathrm{d}S} = \frac{1}{S}$ and as there is no t we have $\frac{\mathrm{d}F}{\mathrm{d}t} = 0$.

The second derivative is $\frac{\mathrm{d}^2F}{\mathrm{d}S^2} = \frac{-1}{S^2}$.

We substitute these values into Itô's lemma given in (3.29) and this gives us

$$\mathrm{d}\ ln\ S = \left(\mu - \frac{\sigma^2}{2}\right)\mathrm{d}t + \sigma\mathrm{d}W. \tag{3.31}$$

So we have moved from dF to dS using Itô's lemma, and (3.31) is a good representation of the asset price over time.

(ii) The bond price equation
The continuously compounded gross redemption yield at time t on a default-free zero-coupon bond that pays £1 at maturity date T is x. We assume that the movement in x is described by

$$\mathrm{d}x = a(\alpha - x)\mathrm{d}t + sx\mathrm{d}Z$$

where a, α and s are positive constants. What is the expression for the process followed by the price P of the bond? Let us say that the price of the bond is given by

$$P = e^{-x(T-t)}.$$

We have dx, and we require dP. This is done by applying Itô's lemma. We require

$$\frac{\partial P}{\partial x} = -(T-t)e^{-x(T-t)} = -(T-t)P$$
$$\frac{\partial^2 P}{\partial x^2} = -(T-t)e^{-x(T-t)} = (T-t)^2 P$$
$$\frac{\partial P}{\partial t} = xe^{-x(T-t)} = xP.$$

From Itô's lemma

$$\mathrm{d}P = \left[\frac{\partial P}{\partial x}a(\alpha - x) + \frac{\partial P}{\partial t} + \frac{1}{2}\frac{\partial^2 P}{\partial x^2}s^2x^2\right]\mathrm{d}t + \frac{\partial P}{\partial x}sx\mathrm{d}Z$$

which gives

$$\mathrm{d}P = \left[-(T-t)Pa(\alpha - x) + xP + \frac{1}{2}(T-t)^2Ps^2x^2\right]\mathrm{d}t - (T-t)Psx\mathrm{d}Z$$

which simplifies to

$$dP = \left[-a(\alpha - x)(T - t) + x + \frac{1}{2}s^2x^2(T - t)^2\right]P\mathrm{d}t - sx(T - t)P\mathrm{d}Z.$$

So using Itô's lemma we have transformed the SDE for the bond yield into an expression for the bond price.

3.2.3 *Stochastic integrals*

Whilst in no way wishing to trivialise the mathematical level, we will not consider the derivations here, but simply state that the observed values of the Brownian motion up to the point at time t determine the process immediately after, and that this process is Gaussian. Stochastic integrals are continuous path martingales. As described in Neftci (1996), the integral is used to calculate sums where we have an infinite or uncountable number of items, in contrast with the Σ sum operator which is used for a finite number of objects. In defining integrals we begin with an approximation, where there is a countable number of items, and then set a limit and move to an uncountable number. A basic definition is given in Appendix 3.4. Stochastic integration is an operation that is closely associated with Brownian paths; a path is partitioned into consecutive intervals or increments, and each increment is multiplied by a random variable. These values are then summed to create the stochastic integral. Therefore the stochastic integral can be viewed as a random walk Brownian motion with increments that have varying values, a random walk with non-homogeneous movement.

3.2.4 *Generalised Itô formula*

It is possible to generalise Itô's formula in order to produce a multi-dimensional formula, which can then be used to construct a model to price interest-rate derivatives or other asset-class options where there is more than one variable. To do this we generalise the formula to apply to situations where the dynamic function $f(\ )$ is dependent on more than one Itô process, each expressed as a standard Brownian motion.

Consider $W_T = \left(W_t^T, \ldots, W_T^n\right)$ where $\left(W_T^t\right)_{t \geq 0}$ are independent standard Brownian motions and W_T is an n-dimensional Brownian motion. We can express Itô's formula mathematically with respect to p Itô processes $\left(X_t^T, \ldots, X_T^t\right)$ as:

$$X_T^i = X_0^i + \int_0^t K_s^i \mathrm{d}s + \sum_{i=1}^{n} \int_0^t H_s^{ij} \mathrm{d}X_s^i. \tag{3.32}$$

Where the function $f(\ )$ contains second-order partial derivatives with respect to x and first-order partial derivatives with respect to t, which are a continuous function in (x, t), the generalised Itô formula is given by

$$\begin{aligned} f(t, X_t^1, \ldots, X_T^p) = f(0, X_0^1, \ldots, X_0^p) &+ \int_0^t \left(\frac{\partial f}{\partial s}\right)(s, X_s^1, \ldots, X_s^p)\mathrm{d}s \\ &+ \sum_{i=1}^{p} \int_0^t \left(\frac{\partial f}{\partial x_i}\right)(s, X_s^1, \ldots, X_s^p)\mathrm{d}X_s^i \\ &+ \frac{1}{2}\sum_{i,j=1}^{p} \int_0^t \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)(s, X_s^1, \ldots, X_s^p)\mathrm{d}(X^i,\ X^j)_s \end{aligned} \tag{3.33}$$

with

$$\begin{aligned} \mathrm{d}X_s^i &= K_s^i \mathrm{d}s + \sum_{j=1}^{n} H_s^{im} \mathrm{d}W_s^j \\ \mathrm{d}(X^i,\ X^j)_s &= \sum_{m=1}^{n} H_s^{im} H_s^{jm} \mathrm{d}s. \end{aligned} \tag{3.34}$$

3.2.5 *Information structures*

A key element of the description of a stochastic process is a specification of the level of information on the behaviour of prices that is available to an observer at each point in time. As with the martingale property, a calculation of the expected future values of a price process requires information on current prices. Generally financial valuation models require data on both the current and the historical security prices, but investors are only able to deal on the basis of current known information, and do not have access to future information. In a stochastic model, this concept is captured via the process known as *filtration*.

A filtration is a family $F = (F_t)$, $t \in T$ of variables $F_t \subset F$ which is increasing in level in the sense that $F_s \subset F_t$ whenever $s,t \in T$, $s \leq t$. Hence a filtration can be viewed as a dynamic information structure, and F_t represents the information available to the investor at time t. The behaviour of the asset price is seen by the increase in filtration, which implies that more and more data is assimilated over time, and historical data is incorporated into the current price, rather than disregarded or forgotten. A filtration $F = (F_t)$ is said to be *augmented* if F_t is augmented for each time t. This means that only F_0 is augmented. A stochastic process W is described as being *adapted* to the filtration F if for each fixed $t \in T$, the random variable X:

$$X_t : w \mapsto X(w,\ t) = X_t(w) = X_t(w) : \Omega \to IR^K$$

is measurable with respect to F_t.

This is an important description as it means that the value X_t of X at t is dependent only on information that is available at time t. It might also mean that an investor with access to the information level F is able to observe or make inferences on the value of X at each point in time. The augmented filtration generated by X is the filtration $F^{\wedge} = (F_t^{\wedge})$, $t \in T$. Any stochastic process X is adapted to the augmented filtration FX that it generates. If a stochastic process is measurable as a mapping or vector then it is a measurable process, however this does not impact significantly on finance theory so we shall ignore it.

3.3 Perfect capital markets

One of the assumptions of derivative pricing is that the financial markets are assumed to be near-perfect, for example akin to Fama's semi-strong or strong-form market. The term *complete* market is also used. Essentially the market is assumed to be a general stochastic economy where transactions may take place at any time, and interest rates behave under *Gaussian uncertainty*. We shall look at this briefly later in this section. Generally pricing models assume that there is an almost infinite number of tradable assets in the market, so that markets are assumed to be complete. This includes the assumptions that there is frictionless continuous trading, with no transaction costs or taxation.

Let us consider then the key assumptions that form part of the economy of for example, the Black–Scholes option pricing model.

3.3.1 *Stochastic price processes*

The uncertainty in asset price dynamics is described as having two sources, both represented by independent standard Brownian motions. These are denoted

$$(W_1t,\ W_2t,\ t \in [0,\ T])$$

on a probability space denoted by (Ω, F, P).

The flow of information to investors is described by the filtration process. The two sources of risk in the Black–Scholes model are the risk-carrying underlying asset, and the cash deposit which, though paying a riskless rate of interest, is at risk from the stochastic character of the interest rate itself.

3.3.2 *Perfect markets*

The assumption of complete capital markets states that, as a result of arbitrage-free pricing, there is a unique probability measure Q, which is identical to the historical probability P, under which the continuously discounted price of any asset is a Q-martingale. This probability level Q then becomes the *risk-neutral* probability.

3.3.3 *Uncertainty of interest rates*

All derivative valuation models describe a process followed by market interest rates. As the future level of the yield curve or spot rate curve is uncertain, the key assumption is that interest rates follow a normal distribution, and follow a Gaussian process. Thus the interest rate is described as being a Gaussian interest rate uncertainty. Only the short-term risk-free interest rate, for which we read the T-bill rate or (in certain situations) the government bond repo rate, is captured in most models. Following Merton (1973), Vasicek (1977), Cox, Ingersoll and Ross (1985) and Jamshidian (1991), the short-dated risk-free interest r applicable to the period t is said to follow a Gaussian diffusion process under a constant volatility. The major drawback under this scenario is that under certain conditions it is possible to model a term structure that produces negative forward interest rates. However in practice this occurs only under certain limited conditions, so the validity of the models is not diminished. The future path followed by r_t is described by the following stochastic differential equation:

$$\mathrm{d}r_t = a_t[b_t - r_t]\mathrm{d}t + \sigma_t \mathrm{d}W_t \tag{3.35}$$

where a and b are constant deterministic functions and σ_t is the instantaneous standard deviation of r_t. Under (3.35) the process describing the returns generated by a risk-free zero-coupon bond $P(t, T)$ that expires at time T and has a maturity of $T-t$ under the risk-neutral probability Q is given by (3.36):

$$\frac{\mathrm{d}P(t,\ T)}{P(t,\ T)} = r_t \mathrm{d}t - \sigma_P(t,\ T)\mathrm{d}W_t \tag{3.36}$$

where σ_P is the standard deviation of the price returns of the $(T-t)$ bond and is a deterministic function defined by:

$$\sigma_P(t,\ T) = \sigma_t \cdot \int_t^T exp\left(-\int_t^u a(s)\mathrm{d}s\right)\mathrm{d}u.$$

In the Black–Scholes model the value of a \$1 (or £1) deposit invested at the risk-free zero-coupon interest rate r and continuously compounded over a period t will have grown to the value given by the expression below, where M_t is the value of the deposit at time t:

$$M_t = exp\left(\int_0^t r(u)\mathrm{d}u\right). \tag{3.37}$$

3.3.4 *Asset price processes*

All valuation models must capture a process describing the dynamics of the asset price. This was discussed at the start of the chapter and is a central tenet of derivatives valuation models. Under the Black–Scholes model for example, the price dynamics of a risk-bearing asset S_t under the risk-neutral probability function Q are given by

$$\frac{\mathrm{d}S_t}{S_t} = r_t\mathrm{d}t + \sigma_S\left(\rho\mathrm{d}W_t + \sqrt{1-\rho^2}\mathrm{d}W_2(t)\right) \tag{3.38}$$

where σ_S is the standard deviation of the asset price returns. The correlation between the price dynamics of the risk-bearing asset and the dynamics of interest-rate changes is given by ρ, $\rho \in [0, 1]$ while $W_2(t)$ is a standard Brownian motion that describes the dynamics of the asset price, and not that of the interest rates which are captured by W_t (and from which it is independent).

Under these four assumptions, the price of an asset can be described in present value terms relative to the value of the risk-free cash deposit M_t and, in fact the price is described as a Q-martingale. A European-style contingent liability with maturity date t is therefore valued at time 0 under the risk-neutral probability as

$$V_0 = E^Q\left[\frac{h_t}{M_t}\right]$$

where

V_0	is the value of the asset at time 0
h_t	is the stochastic payoff at maturity date t, where h is a measurable stochastic process
and $E^Q[\ ]$	is the expectation of the value under probability function Q.

In the following chapter we tie in the work on dynamics of asset prices to option valuation models.

Appendices

Appendix 3.1: An introduction to stochastic processes

A stochastic process can be described with respect to the notion of a vector of variables. If we set the following parameters

Ω is the set of all possible states ε

ψ is a class of partitions of Ω

$X(\omega)$ is said to be a *random variable* when it is a measurable application from (Ω, ψ) to $\Re$. A vector of random variables $X(\omega) = [X_1(\omega), \ldots, X_n(\omega)]$ is an application that can be measured from (Ω, ψ) into $\Re^n$. Therefore we have a vector of random variables that is similar to n ordinary variables defined under the same probability function.

A stochastic process is an extension of the notion of a vector of variables when the number of elements becomes infinite. It is described by

$$\{X_t(\omega)\},\ t \in T$$

which is a set of random variables where the index varies in a finite or infinite group, and is denoted by $X(t)$.

Appendix 3.2: Itô's lemma

If f is a continuous and differentiable function of a variable x, and Δx is a small change in x, then using a Taylor expansion the resulting change in f is given by (3.39):

$$\Delta f = \left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)\Delta x + \frac{1}{2}\left(\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\right)\Delta x^2 + \frac{1}{6}\left(\frac{\mathrm{d}^3 f}{\mathrm{d}x^3}\right)\Delta x^3 + \cdots. \tag{3.39}$$

If f is dependent on two variables x and y then the Taylor expansion of Δf becomes (3.40):

$$\Delta f \approx \left(\frac{\partial f}{\partial x}\right)\Delta x + \left(\frac{\partial f}{\partial y}\right)\Delta y + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)\Delta x^2 + \frac{1}{2}\left(\frac{\partial^2 f}{\partial y^2}\right)\Delta y^2 + \left(\frac{\partial^2 f}{\partial x \partial y}\right)\Delta x \Delta y + \cdots. \tag{3.40}$$

The limiting case where Δx and Δy are close to zero will transform (3.40) to (3.41):

$$\partial f \approx \left(\frac{\partial f}{\partial x}\right)\mathrm{d}x + \left(\frac{\partial f}{\partial y}\right)\mathrm{d}y. \tag{3.41}$$

Consider now a derivative asset $f(x, \tau)$ whose value is dependent on time and on the asset price x. If we assume that x follows the general Itô process

$$\mathrm{d}x = a(x,\ t)\mathrm{d}t + b(x,\ t)\mathrm{d}W \tag{3.42}$$

where a and b are functions of x and t and dW is a Weiner process. The asset price x is described by a drift rate of x and a standard deviation of b. Using Itô's lemma it can be shown that a function f of x and t will follow the following process:

$$\mathrm{d}f = \left(\frac{\partial f}{\partial x}a + \frac{\partial f}{\partial t} + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)b^2\right)\mathrm{d}t + \frac{\partial f}{\partial x}b\mathrm{d}W \tag{3.43}$$

and where dW is the Weiner process; therefore f follows an Itô process and its drift and standard deviation are described by the expressions below:

$$\frac{\partial f}{\partial x}a+\frac{\partial f}{\partial t}+\frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)b^2 \qquad \text{and} \qquad \left(\frac{\partial f}{\partial x}\right)b.$$

This may also be stated as:

$$\Delta x = a(x,t)\Delta t + b\varepsilon\sqrt{\Delta t}$$

where the term ε is normally distributed with a mean of 0, so that $E(\varepsilon)=0$, and a variance of 1, so that $E(\varepsilon^2)-E(\varepsilon)^2=1$. In the limit case (3.40) becomes (3.44), which is Itô's lemma:

$$\mathrm{d}f=\left(\frac{\partial f}{\partial x}\right)\mathrm{d}x+\left(\frac{\partial f}{\partial t}\right)\mathrm{d}t+\frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)b^2\mathrm{d}t. \tag{3.44}$$

The expression in (3.44) is Itô's lemma and if we substitute (3.42) for dx, it can be transformed to (3.45):

$$\mathrm{d}f=\left[\left(\frac{\partial f}{\partial x}\right)a+\left(\frac{\partial f}{\partial t}\right)+\frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)b^2\right]\mathrm{d}t+\left(\frac{\partial f}{\partial x}\right)b\mathrm{d}W. \tag{3.45}$$

The derivation of Itô's formula is given in Appendix 3.3.

Appendix 3.3: Derivation of Itô's formula

Let X_t be a *stochastic* process described by

$$\mathrm{d}X_t=\mu_t\mathrm{d}t+\sigma_t\mathrm{d}W_t \tag{3.46}$$

where W_t is a random variable and Brownian motion and dW_t is an incremental change in the Brownian motion W_t, equal to $Z_t\sqrt{\mathrm{d}t}$, $Z_t\sim N(0,\,1)$. Then suppose that we have a function $Y_t=f(X_t,\,t)$ and we require the differential dY_t. Applying a Taylor expansion of Y_t we would obtain

$$\mathrm{d}Y_t=\frac{\partial f}{\partial X_t}\mathrm{d}X_t+\frac{\partial f}{\partial t}\mathrm{d}t+\frac{1}{2}\left[\frac{\partial^2 f}{\partial X_t^2}\mathrm{d}X_t^2+2\frac{\partial^2 f}{\partial X\partial t}\mathrm{d}X_t\mathrm{d}t+\frac{\partial^2 f}{\partial t^2}\mathrm{d}t^2\right]+\cdots. \tag{3.47}$$

In (3.46) if we square dX_t we obtain

$$\mathrm{d}X_t^2=\mu_t^2\mathrm{d}t^2+2\sigma_t\mu_t\mathrm{d}W_t\mathrm{d}t+\sigma_t^2\mathrm{d}W_t^2. \tag{3.48}$$

The first two terms in (3.48) are of a higher order and of minimal impact when dt is sufficiently small, and may be ignored. It can be shown that the variance of the $(\mathrm{d}W_t)^2$ term will tend towards zero when the increment dt is sufficiently small. At this point it no longer has the property of a random variable and becomes more a constant with expected value

$$E(Z^2\mathrm{d}t)=\mathrm{d}t. \tag{3.49}$$

It can then be shown that for sufficiently small dt

$$\mathrm{d}X_t^2\Rightarrow\sigma_t^2\mathrm{d}W_t^2\Rightarrow\sigma_t^2\mathrm{d}t.$$

The differential $\mathrm{d}Y_t$ has an element that tends towards $\frac{1}{2}\frac{\partial^2 f}{\partial X_t^2}\sigma_t^2\mathrm{d}t$ for sufficiently small $\mathrm{d}t$ but cannot be dropped as were the higher-order terms of (3.48) as it is of order $\mathrm{d}t$. So the first-order differential of Y_t is

$$\mathrm{d}Y_t = \frac{\partial f}{\partial X_t}\mathrm{d}X_t + \frac{\partial f}{\partial t}\mathrm{d}t + \frac{1}{2}\sigma_t^2\frac{\partial^2 f}{\partial X_t^2}\mathrm{d}t \tag{3.50}$$

and now if we insert $\mathrm{d}X_t$ from (3.46) into (3.49) we will obtain

$$\mathrm{d}Y_t = \left(\mu_t\frac{\partial f}{\partial X_t} + \frac{\partial f}{\partial t} + \frac{1}{2}\sigma_t^2\frac{\partial^2 f}{\partial X_t^2}\right)\mathrm{d}t + \sigma_t\frac{\partial f}{\partial X_t}\mathrm{d}W_t. \tag{3.51}$$

If the reader has followed this through he or she has arrived at Itô's lemma. We can apply this immediately. Consider a process

$$X_t = X_0 + \int_0^t \mu_\nu \mathrm{d}\nu + \int_0^t \sigma_s \mathrm{d}W_s \tag{3.52}$$

for which the differential form is

$$\mathrm{d}X_t = \mu_t\mathrm{d}t + \sigma_t\mathrm{d}W_t. \tag{3.53}$$

If we set the function $f(X)$ equal to X_t the results of applying the Itô lemma terms are

$$\frac{\partial f}{\partial X_t} = 1; \quad \frac{\partial f}{\partial t} = 0 \quad \text{and} \quad \frac{\partial^2 f}{\partial X_t^2} = 0.$$

Therefore using Itô's lemma we obtain

$$\mathrm{d}X_t = \mu_t\mathrm{d}t + \sigma_t\mathrm{d}W_t \tag{3.54}$$

which is what we expect. What we have here is a stochastic differential equation at (3.54) for which the solution is (3.52).

Appendix 3.4: The integral

Suppose we have a deterministic function $f(x)$ of time, with $x \in [0, T]$ that corresponds to a curve of $f(x)$ over the period from 0 to T, and we wish to calculate the area given by the function from time t_0 to t_T. This can be done by integrating the function over the time interval $[0, T]$, given by

$$\int_0^T f(s)\mathrm{d}s. \tag{3.55}$$

To calculate the integral we split the area given by the function in the time period into a series of *partitions* or intervals, described by

$$t_0 = 0 < t_1 < t_2 < \ldots < t_n = T. \tag{3.56}$$

The approximate value of the area required is

$$\sum_{i=1}^{n} f\left(\frac{t_i + t_{i-1}}{2}\right)(t_i - t_{i-1}); \tag{3.57}$$

however, if we decrease the interval space such that it approaches 0, described by

$$\max_i \left| t_i - t_{i-1} \right| \to 0$$

the area under the space is given by the integral in (3.49), as the approximating sum approaches the area defined by the limit

$$\sum_{i=1}^{n} f\left(\frac{t_i + t_{i-1}}{2}\right)(t_i - t_{i-1}) \to \int_0^T f(s)\mathrm{d}s. \tag{3.58}$$

Selected bibliography and references

Baxter, M., Rennie, A., *Financial Calculus*, Cambridge University Press, 1996.

Boyle, P., 'Option valuation using a three jump process', *International Options Journal* 3, 1986, pp. 7–12.

Briys, E., Bellalah, M., Mai, H., de Varenne, F., *Options, Futures and Exotic Derivatives*, Wiley, 1998.

Cootner, P. (ed.), *The Random Character of Stock Market Prices*, MIT Press, 1964.

Cox, D., Miller, H., *The Theory of Stochastic Processes*, Chapman & Hall, 1965.

Cox, J., Ingersoll, J., Ross, S., 'A theory of the term structure of interest rates', *Econometrica* 53, 1985, pp. 385–407.

Debreu, G., 'Representation of a preference ordering by a numerical function', in Thrall, R., Coombs, C., Davis, R. (eds.), *Decision Processes*, Wiley, 1954.

Duffie, D., *Dynamic Asset Pricing Theory*, Princeton University Press, 1996.

Fama, E., 'The behaviour of stock prices', *Journal of Business* 38, January 1965, pp. 34–105.

Feller, W., *Probability Theory and Its Applications*, vols. 1 and 2, Wiley, 1950.

Harrison, J., Kreps, D., 'Martingales and arbitrage in multi-period securities markets', *Journal of Economic Theory* 20, 1979, pp. 381–408.

Harrison, J., Pliska, S., 'Martingales and stochastic integrals in the theory of continuous trading', *Stochastic Processes and Their Applications* 11, 1981, pp. 216–260.

Hull, J., *Options, Futures and Other Derivatives*, Wiley, 1997.

Ingersoll, J., Jr, *Theory of Financial Decision Making*, Rowman & Littlefield, 1987.

Itô, K., 'On stochastic differential equations', *American Mathematical Society* 4, 1951, pp. 1–51.

James, J., Webber, N., *Interest Rate Modelling*, Wiley 2000, Chapters 3–5, 7–9, 15–16.

Jamshidian, F., 'Bond and option valuation in the Gaussian interest rate model', *Research in Finance* 9, 1991, pp. 131–170.

Merton, R., 'Theory of rational option pricing', *Bell Journal of Economics and Management Science* 4, Spring 1973, pp. 141–183.

Neftci, S., *An Introduction to the Mathematics of Financial Derivatives*, Academic Press, 1996.

Nielson, L.T., *Pricing and Hedging of Derivative Securities*, Oxford University Press, 1999.

Rendleman, R., Bartter, B., 'Two state option pricing', *Journal of Finance* 34, 1979, pp. 1092–1110.

Ross, S., *Probability Models*, 7th edition, Harcourt Academic Press, 2000.

Vasicek, O., 'An equilibrium characterisation of the term structure', *Journal of Financial Economics* 5, 1977, pp. 177–188.

Williams, D., *Probability with Martingales*, Cambridge University Press, 1991.

4 Interest-rate Models I

In Chapter 3 we introduced the concept of stochastic processes. Most but not all interest rate models are essentially a description of the short-rate in terms of a stochastic process. Recent literature[1] has tended to categorise models into one of up to six different types, but for our purposes we can generalise them into two types. Thus we introduce some of the main models, according to their categorisation as equilibrium or arbitrage-free models. This chapter looks at the earlier models, including the first ever term structure model presented by Vasicek in 1977. The next chapter considers what have been termed 'whole yield curve' models, or the Heath–Jarrow–Morton family, while Chapter 6 reviews considerations in fitting the yield curve.

4.1 Introduction

4.1.1 *Bond price and yield*

We first set the scene by introducing the interest rate market. The price of a zero-coupon bond of maturity T at time t is denoted by $P(t, T)$, so that its price at time 0 is denoted by $P(0, T)$. The process followed by the bond price is a stochastic one and therefore can be modelled, equally options that have been written on the bond can be hedged by it. If market interest rates are constant, the price of the bond at time t is given by $e^{-r(T-t)}$. This enables us to state that given a zero-coupon bond price $P(t, T)$ at time t, the yield $r(t, T)$ is given by (4.1):

$$r(t,\ T) = -\frac{log\ P(t,\ T)}{T - t}. \tag{4.1}$$

Of course interest rates are not constant but (4.1) is valuable as it is used later in constructing a model. By using (4.1) we are able to produce a yield curve given a set of zero-coupon bond prices. For modelling purposes we require a definition of the *short rate*, or the current interest rate for borrowing a sum of money that is paid back a very short period later (in fact, almost instantaneously). This is the rate payable at time t for repayment at time $t + \Delta t$ where Δt is an incremental passage of time. This is given by

$$r(t,\ t + \Delta t) = -\frac{log\ P(t,\ t + \Delta t)}{\Delta t} \tag{4.2}$$

and the incremental change can be steadily decreased to give the instantaneous rate, which is described by

$$r_t = -\frac{\partial}{\partial T} log\ P(t,\ t) \tag{4.3}$$

and is identical to $r(t, t)$.

---

1 For example, see James and Webber (2000), or Van Deventer and Imai (1997).

The instantaneous rate is an important mathematical construct that is widely used in the modelling process.

We can define forward rates in terms of the short rate. Again for infinitesimal change in time from a forward date T_1 to T_2 (for example, two bonds whose maturity dates are very close together), we can define a forward rate for instantaneous borrowing, given by

$$rf(t,\ T) = -\frac{\partial}{\partial T} log\ P(t,\ T) \tag{4.4}$$

which is called the forward rate. We can also set

$$rf(t,\ t) = r_t \tag{4.5}$$

that is the forward rate for borrowing at the point $t = T$ which is identical to the short rate. The forward rate is valuable because, given the set of forward rates from t to T, we can calculate the bond price for a T-maturity date. This is presented in a number of texts, one of the best being Jarrow (1996). Given the expressions for the bond yield and the forward rate, the bond prices can be defined in terms of either the yield,

$$P(t,\ T) = exp(-(T-t)r(t,\ T)) \tag{4.6}$$

or the forward rates, as a stochastic integral

$$P(t,\ T) = exp\left(-\int_t^T rf(t,\ s)\mathrm{d}s\right). \tag{4.7}$$

This is convenient because this means that the price at time t of a zero-coupon bond maturing at T is given by (4.7), and forward rates can be calculated from the current term structure or vice-versa.

For readers unfamiliar with the basic maths, an introductory primer is given in the Preface, which can be used to reference the relevant texts.

4.1.2 *Interest rate models*

An interest rate model provides a description of the dynamic process by which rates change over time, in terms of a statistical construct, as well as a means by which interest rate derivatives such as options can be priced. It is often the practical implementation of the model that dictates which type is used, rather than mathematical neatness or more realistic assumptions. An excellent categorisation is given in James and Webber (2000) who list models as being one of the following types:

- the traditional one-, two- and multi-factor equilibrium models, known as *affine term structure* models (see James and Webber (2000) or Duffie (1996, p. 136)).[2] These include Gaussian affine models such as Vasicek, Hull–White and Steeley, where the model describes a process with constant volatility; and models that have a square-root volatility such as Cox–Ingersoll–Ross (CIR);
- whole yield curve models such as Heath–Jarrow–Morton;

[2] A function $\Re \rightarrow \Re$ is *affine* if there are constants a and b such that for all values of x, $H(x) = a + bx$. This describes certain term structure models' drift and diffusion functions.

- so-called market models such as Jamshidian;
- so-called consol models such as Brennan and Schwartz.

There are also other types of models and we suggest that interested readers consult a specialist text; James and Webber is an excellent start, which also contains detailed sections on implementing models as well as a comparison of the different models themselves.

The most commonly used models are the Hull–White type models which are relatively straightforward to implement, although HJM models are also more commonly encountered. The Hull–White and *extended* CIR models incorporate a mean reversion feature that means that they can be fitted to the term structure in place at the time. The CIR model has a square-root factor in its volatility component, which prevents the short-term rate reaching negative values. What criteria are used by a bank in deciding which model to implement? Generally a user will seek to implement a model that fits current market data, fits the process by which interest rates change over time and is *tractable*. This means that it should be computationally efficient, and provide explicit solutions when used for pricing bonds and vanilla options.

4.1.3 *Introduction to bond analysis using spot rates and forward rates in continuous time*

This section analyses further the relationship between spot and forward rates and the yield curve.

The spot and forward rate relationship

In the discussion to date, we have assumed discrete time intervals and interest rates in discrete time. Here we consider the relationship between spot and forward rates in continuous time. For this we assume the mathematical convenience of a continuously compounded interest rate.

The rate r is compounded using e^r and an initial investment M earning $r(t, T)$ over the period $T-t$, initial investment at time t and for maturity at T, where $T > t$, would have a value of $Me^{r(t,\,T)(T-t)}$ on maturity.[3] If we denote the initial value M_t and the maturity value M_T then we can state $M_t e^{r(t,\,T)(T-t)} = M_T$ and therefore the continuously compounded yield, defined as the continuously compounded interest rate $r(t, T)$ can be shown to be

$$r(t,\ T) = \frac{log(M_T/M_t)}{T-t}. \tag{4.8}$$

We can then formulate a relationship between the continuously compounded interest rate and yield. It can be shown that

$$M_T = M_t e^{\int_t^T r(s)\mathrm{d}s} \tag{4.9}$$

where $r(s)$ is the instantaneous spot interest rate and is a function of time. It can further be shown that the continuously compounded yield is actually the equivalent of the average value of the continuously compounded interest rate. In addition it can be shown that

$$r(t,\ T) = \frac{\int_t^T r(s)\mathrm{d}s}{T-t}. \tag{4.10}$$

---

[3] e is the mathematical constant 2.7182818… and it can be shown that an investment of £1 at time t will have grown to e on maturity at time T (during the period $T-t$) if it is earning an interest rate of $1/(T-t)$ continuously compounded.

In a continuous time environment we do not assume discrete time intervals over which interest rates are applicable, rather a period of time in which a borrowing of funds would be repaid instantaneously. So we define the forward rate $f(t, s)$ as the interest rate applicable for borrowing funds where the deal is struck at time t; the actual loan is made at s (with $s > t$) and repayable almost instantly. In mathematics the period $s-t$ is described as infinitesimally small. The spot interest rate is defined as the continuously compounded yield or interest rate $r(t, T)$. In an environment of no arbitrage, the return generated by investing at the forward rate $f(t, s)$ over the period $s-t$ must be equal to that generated by investing initially at the spot rate $r(t, T)$. So we may set

$$e^{\int_t^T f(t,\ s)\mathrm{d}s} = e^{r(t)\mathrm{d}t} \tag{4.11}$$

which enables us to derive an expression for the spot rate itself, which is

$$r(t,\ T) = \frac{\int_t^T f(t,\ s)\mathrm{d}s}{T-t}. \tag{4.12}$$

The relationship described by (4.12) states that the spot rate is given by the *arithmetic* average of the forward rates $f(t, s)$ where $t < s < T$. How does this differ from the relationship in a discrete time environment? We know that the spot rate in such a framework is the *geometric* average of the forward rates,[4] and this is the key difference in introducing the continuous time structure. Equation (4.12) can be rearranged to

$$r(t,\ T)(T-t) = \int_t^T f(t,\ s)\mathrm{d}s \tag{4.13}$$

and this is used to show (by differentiation) the relationship between spot and forward rates, given below:

$$f(t,\ s) = r(t,\ T) + (T-t)\frac{\mathrm{d}r(t,\ T)}{\mathrm{d}T}. \tag{4.14}$$

If we assume we are dealing today (at time 0) for maturity at time T, then the expression for the spot rate becomes

$$r(0,\ T) = \frac{\int_0^T f(0,\ s)\mathrm{d}s}{T} \tag{4.15}$$

so we can write

$$r(0,\ T)\cdot T = \int_0^T f(0,\ s)\mathrm{d}s. \tag{4.16}$$

This is illustrated in Figure 4.1 which is a diagrammatic representation showing that the spot rate $r(0, T)$ is the average of the forward rates from 0 to T, using the hypothetical value of 5% for $r(0, T)$. Figure 4.1 also shows the area represented by (4.16).

[4] To be precise, if we assume annual compounding, the relationship one plus the spot rate is equal to the geometric average of one plus the forward rates.

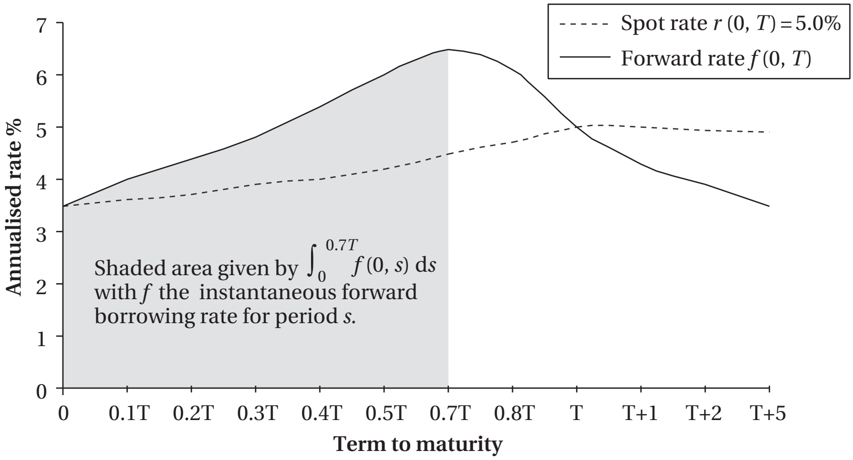


Figure 4.1: Diagrammatic representation of the relationship between spot and forward rate. The spot rate $r(t, T)$ is the average of the forward rates between t and T.

What (4.14) implies is that if the spot rate increases, then by definition the forward rate (or *marginal* rate as has been suggested that it may be called[5]) will be greater. From (4.14) we deduce that the forward rate will be equal to the spot rate plus a value that is the product of the *rate* of increase of the spot rate and the time period $(T - t)$. In fact the conclusions simply confirm that the forward rate for any period will lie above the spot rate if the spot rate term structure is increasing, and will lie below the spot rate if it is decreasing. In a constant spot rate environment, the forward rate will be equal to the spot rate.

However it is not as simple as that. An increasing spot rate term structure only implies that the forward rate lies above the spot rate, but not that the forward rate structure is itself also *increasing*. In fact one can observe the forward rate term structure to be increasing or decreasing while spot rates are increasing. As the spot rate is the average of the forward rates, it can be shown that in order to accommodate this, forward rates must in fact be *decreasing* before the point at which the spot rate reaches its highest point. This confirms market observation. An illustration of this property is given in Appendix 4.1. As Campbell *et al.* (1997) state, this is a property of average and marginal cost curves in economics.

5 For example see Section 10.1 of Campbell, Lo and MacKinlay (1997), Chapter 10 of which is an excellent and accessible study of the term structure, and provides proofs of some of the results discussed here. This book is written in very readable style and is worth purchasing for Chapter 10 alone.

Bond prices as a function of spot and forward rates

In this section we describe the relationship between the price of a zero-coupon bond and spot and forward rates. We assume a risk-free zero-coupon bond of nominal value £1, priced at time t and maturing at time T. We also assume a money market bank account of initial value $P(t, T)$ invested at time t. The money market account is denoted M. The price of the bond at time t is denoted $P(t, T)$ and if today is time 0 (so that $t > 0$) then the bond price today is unknown and a random factor (similar to a future interest rate). The bond price can be related to the spot rate or forward rate that is in force at time t.

Consider the scenario below, used to derive the risk-free zero-coupon bond price.[6]

The continuously compounded *constant* spot rate is r as before. An investor has a choice of purchasing the zero-coupon bond at price $P(t, T)$, which will return the sum of £1 at time T or of investing this same amount of cash in the money market account, and this sum would have grown to £1 at time T. We know that the value of the money market account is given by $Me^{r(t,\,T)(T-t)}$. If M must have a value of £1 at time T then the function $e^{-r(t,\,T)(T-t)}$ must give the present value of £1 at time t and therefore the value of the zero-coupon bond is given by

$$P(t,\ T) = e^{-r(t,\,T)(T-t)}. \tag{4.17}$$

If the same amount of cash that could be used to buy the bond at t, invested in the money market account, does *not* return £1 then arbitrage opportunities will result. If the price of the bond exceeded the discount function $e^{-r(t,\,T)(T-t)}$ then the investor could short the bond and invest the proceeds in the money market account. At time T the bond position would result in a cash outflow of £1, while the money market account would be worth £1. However the investor would gain because in the first place $P(t, T) - e^{-r(t,\,T)(T-t)} > 0$. Equally if the price of the bond was below $e^{-r(t,\,T)(T-t)}$ then the investor would borrow $e^{-r(t,\,T)(T-t)}$ in cash and buy the bond at price $P(t, T)$. On maturity the bond would return £1, which proceeds would be used to repay the loan. However the investor would gain because $e^{-r(t,\,T)(T-t)} - P(t,\ T) > 0$. To avoid arbitrage opportunities we must therefore have

$$P(t,\ T) = e^{-r(t,\,T)(T-t)}. \tag{4.18}$$

Following the relationship between spot and forward rates it is also possible to describe the bond price in terms of forward rates.[7] We show the result here only. First we know that

$$P(t,\ T)e^{\int_t^T f(t,\,s)\mathrm{d}s} = 1 \tag{4.19}$$

because the maturity value of the bond is £1, and we can rearrange (4.19) to give

$$P(t,\ T) = e^{-\int_t^T f(t,\,s)\mathrm{d}s}. \tag{4.20}$$

Expression (4.20) states that the bond price is a function of the range of forward rates that apply for all $f(t, s)$ that is, the forward rates for all time periods s from t to T (where $t < s < T$, and where s is infinitesimally small). The forward rate $f(t, s)$ that results for each s arises as a

6 This approach is also used in Campbell *et al.* (q.v.).

7 For instance, see *ibid*, Section 4.2.

result of a random or *stochastic* process that is assumed to start today at time 0. Therefore the bond price $P(t, T)$ also results from a random process, in this case all the random processes for all the forward rates $f(t, s)$.

The zero-coupon bond price may also be given in terms of the spot rate $r(t, T)$, as shown in (4.18). From our earlier analysis we know that

$$P(t, T)e^{r(t, T)(T-t)} = 1 \tag{4.21}$$

which is rearranged to give the zero-coupon bond price equation

$$P(t, T) = e^{-r(t, T)(T-t)} \tag{4.22}$$

as before.

Equation (4.22) describes the bond price as a function of the spot rate only, as opposed to the multiple processes that apply for all the forward rates from t to T. As the bond has a nominal value of £1 the value given by (4.22) is the discount factor for that term; the range of zero-coupon bond prices would give us the discount function.

What is the importance of this result for our understanding of the term structure of interest rates? First, we see (again, but this time in continuous time) that spot rates, forward rates and the discount function are all closely related, and given one we can calculate the remaining two. More significantly, we may model the term structure either as a function of the spot rate only, described as a stochastic process, or as a function of all of the forward rates $f(t, s)$ for each period s in the period $(T - t)$, described by multiple random processes. The first yield curve models adopted the first approach, while a later development described the second approach.

4.2 Interest-rate processes

Term structure models are essentially models of the interest-rate process. The problem being posed is, what behaviour is exhibited by interest rates, and by the short-term interest rate in particular? An excellent description of the three most common processes that are used to describe the dynamics of the short-rate is given in Phoa (1998), who describes:

- **the Gaussian or normal process:** random shifts in forward rates are normally distributed and any given forward rate drifts upward at a rate proportional to the initial time to the forward date. The interest-rate volatility is independent of the current interest rate, and the volatility term has the form $\sigma \mathrm{d}W(t)$ where $W(t)$ is a generalised Weiner process or Brownian motion. An example of a Gaussian model is the Vasicek model;
- **the square root or squared Gaussian process:** the interest-rate volatility is proportional to the square root of the current interest rate, so the volatility term is given by $\sigma\sqrt{r}\mathrm{d}W(t)$. An example of this is the Cox–Ingersoll–Ross model;
- **the lognormal process:** interest-rate volatility is proportional to the current interest rate, with the volatility term described by $\sigma r \mathrm{d}W(t)$. An example of this is the Black–Derman–Toy model.

To illustrate the differences, this means that if the current short-rate is 8% and is assumed to have an annualised volatility of 100 basis points, and at some point in the future the short-rate moves to 4%, under the Gaussian process the volatility at the new rate will remain at 50 basis points, the square root process will assume a volatility of 82.8 basis points and the lognormal process will assume a volatility of 50 basis points.

The most straightforward models to implement are normal models, followed by square root models and then lognormal models. The process that is used will have an impact on the distribution of future interest rates predicted by the model. A generalised distribution is given in Figure 4.2.

Empirical studies have not pointed conclusively to one specific process as the most realistic. One study (BoE 1999) states that observation of interest rate behaviour in different markets suggests that when current interest rate levels are low, at 4% or below, the rate process has tended to a Gaussian process, while when rates are relatively high the process is more akin to a lognormal process. At levels between these two, it would seem an 'intermediate' process is followed. These observations can be supported by economic argument however. The nominal level of interest rates in an economy has two elements, a real interest rate and an inflation component. Thus interest-rate volatility arises as a result of real interest-rate volatility and consumer prices volatility. When interest rates are low, the inflation component will be negligible, at which point only real rate volatility has an impact. However as real rates are linked to the rate of growth, it is reasonable to assume that they follow a normal distribution. An extreme case has occurred in some markets where the real rates on index-linked bonds have occasionally been recorded as negative. When interest rates are at relatively high levels, the inflation component is more significant, so that price volatility is important. However economic rationale suggests that the price of traded goods follows a lognormal distribution.

Where does this leave the thinking on interest-rate models? As we demonstrate in the next section, one of the drawbacks of Gaussian interest-rate models is that they can result in negative forward rates. Although not impossible, this is an extremely unusual, not to say rare, situation and one that is unlikely in any environment bar one with very low current

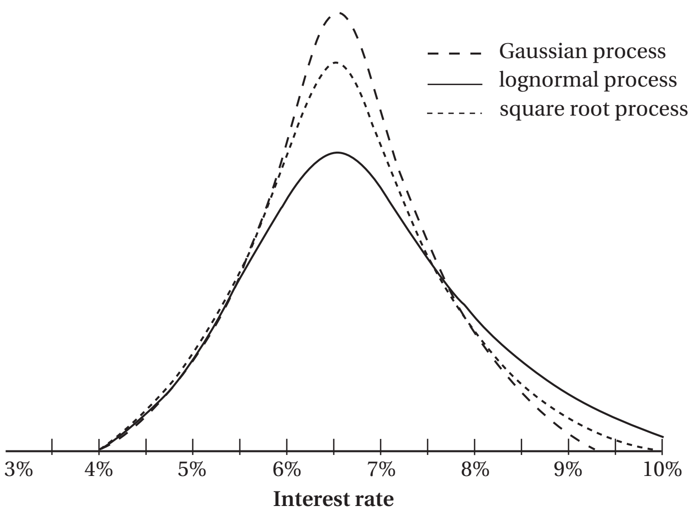


Figure 4.2: Distribution of future interest rates implied by different processes.

interest rates. However such a phenomenon is not completely unheard of, and an environment of low interest rates is one that is best described by a Gaussian process. Negative interest rates have been recorded, for example in the Japanese government bond repo market and certain other repo markets when bonds have gone very *special*, and bear in mind that rates in Japan have been very low for some time now. Essentially then a model that permits negative interest rates is not necessarily unrealistic in an economic sense.

4.3 One-factor models

A short-rate model can be used to derive a complete term structure. We can illustrate this by showing how the model can be used to price discount bonds of any maturity. The derivation is not shown here. Let $P(t, T)$ be the price of a risk-free zero-coupon bond at time t maturing at time T that has a maturity value of 1. This price is a random process, although we know that the price at time T will be 1. Assume that an investor holds this bond, which has been financed by borrowing funds of value C_t. Therefore at any time t the value of the short cash position must be $C_t = -P(t, T)$, otherwise there would be an arbitrage position. The value of the short cash position is growing at a rate dictated by the short-term risk-free rate r, and this rate is given by

$$\frac{\mathrm{d}C_t}{\mathrm{d}t} = r(t)C_t.$$

By integrating this we obtain $C_t = C_0 \; exp(-\int_0^t r(s)\mathrm{d}s)$ which can be rearranged to give

$$P(0, T)/P(t, T) = exp\left(-\int_0^t r(s)\mathrm{d}s\right)$$

so that the random process on both sides is the same, so that their expected values are the same. This can be used to show that the price of the zero-coupon bond at any point t is given by:

$$P(t, T) = E\left[exp\left(-\int_t^T r(s)\mathrm{d}s\right)\right].$$

Therefore, once we have a full description of the random behaviour of the short-rate r, we can calculate the price and yield of any zero-coupon bond at any time, by calculating this expected value. The implication is clear: specifying the process $r(t)$ determines the behaviour of the entire term structure, so if we wish to build a term structure model we need only (under these assumptions) specify the process for $r(t)$.

So now we have determined that a short-rate model is related to the dynamics of bond yields and therefore may be used to derive a complete term structure. We also said that in the same way the model can be used to value bonds of any maturity. The original models were one-factor models, which describe the process for the short-rate r in terms of one source of uncertainty. This is used to capture the short-rate in the following form:

$$\mathrm{d}r = \mu(r)\mathrm{d}t + \sigma(r)\mathrm{d}W \tag{4.23}$$

where μ is the instantaneous drift rate and σ the standard deviation of the short-rate r. Both these terms are assumed to be functions of the short-rate and independent over time. The key assumption made in a one-factor model is that all interest rates move in the same direction.

4.3.1 *The Vasicek model*

In the Vasicek model (1977) the instantaneous short-rate r is assumed to follow a stochastic process known as the Ornstein–Uhlenbeck process, a form of Gaussian process, described by (4.24):

$$dr = a(b - r)dt + \sigma dW. \tag{4.24}$$

This model incorporates *mean reversion*, which is a not unrealistic feature. Mean reversion is the process that describes that when the short-rate r is high, it will tend to be pulled back towards the long-term average level; when the rate is low, it will have an upward drift towards the average level. In Vasicek's model the short-rate is pulled to a mean level b at a rate of a. The mean reversion is governed by the stochastic term σdW which is normally distributed. Using (4.24) Vasicek shows that the price at time t of a zero-coupon bond of maturity T is given by:

$$P(t,\ T) = A(t,\ T)e^{-B(t,\ T)r(t)} \tag{4.25}$$

where $r(t)$ is the value of r at time t,

$$B(T,\ t) = \frac{1 - e^{-A(t-T)}}{a} \tag{4.26}$$

$$\text{and} \quad A(t,\ T) = exp\left(\frac{B(t,\ T) - (T - t)(a^2 b - (\sigma^2/2))}{a^2} - \frac{\sigma^2 B(t,\ T)^2}{4a}\right). \tag{4.27}$$

It can be shown further that

$$r(t,\ T) = -\frac{1}{T-t} ln\ A(t,\ T) + \frac{1}{T-t} B(t,\ T) r(T) \tag{4.28}$$

which describes the complete term structure as a function of $r(t)$ with parameters a, b and the standard deviation σ. The expression in (4.28) states that $r(t, T)$ is a linear function of $r(t)$, and that the value of $r(t)$ will determine the level of the term structure at time t. Using the parameters described above we can calculate the price function for a risk-free zero-coupon bond. Chan *et al.* (1992) used the following parameters: a long-run mean b of 0.07, drift rate a of 0.18 and standard deviation of 0.02. Using these parameters Figure 4.3 shows two zero-

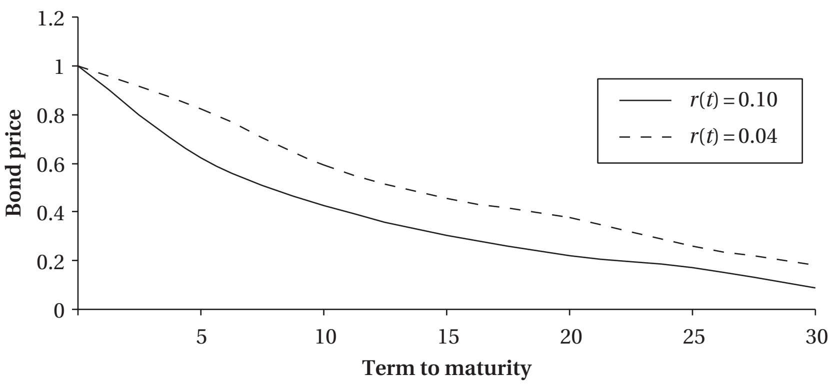


Figure 4.3: Zero-coupon bond price curves at $r(t) = 0.04$ and $r(t) = 0.10$.

coupon bond price curves that result from two different initial short rates, $r(t) = 4\%$ and $r(t) = 10\%$. The time to maturity T is measured on the x-axis, with the price of the zero-coupon bond with that time to maturity (a redemption value of 1) measured along the y-axis.

For derived forward rates the bond price function $P(t, T)$ continuously differentiable with respect to t. Therefore the model produces the following for the instantaneous forward rates:

$$\begin{aligned} f(t,\ T) &= -\frac{\partial}{\partial t} ln\ P(t,\ T) \\ &= A'(T-t) + B'(T-t)r(t) = \left(1 - e^{-a(T-t)}\right)b - \frac{1}{2}\nu(T-t) + e^{-a(T-t)r(t)} \\ &= f(r(t),\ T-t) \end{aligned} \tag{4.29}$$

where $f(r, T)$ is the function $f(r, T) = (1 - e^{-aT})b - \frac{1}{2}\nu(T) + e^{-aT}r$.

The forward rate is a function of the short-rate and is normally distributed. Figure 4.4 shows the forward rate curves that correspond to the price curves in Figure 4.3, under the same parameters.

An increase in the initial short-rate r will have the effect of raising forward rates, as will increasing the long-run mean value b. The effect of an increase in r is most pronounced at shorter maturities, whereas an increase in b has the greatest effect the longer the term to maturity. An equal increase or decrease in both r and b will have the effect of moving all forward rates up or down by the same amount. With these changes the forward curve moves up or down in a parallel fashion.

The derived forward rate is a decreasing function of the instantaneous standard deviation σ, one of the model parameters. The partial derivative of the forward rate with respect to the standard deviation is given in (4.30):

$$\frac{\partial f(r,\ T)}{\partial \sigma} = -\sigma B(T-s)^2 = -\frac{\sigma}{a^2}(-2e^{-aT} + e^{-2aT} + 1). \tag{4.30}$$

The expression in (4.30) states that the forward rate is a decreasing function of T, that is it becomes more negative as T becomes larger. The effect of the standard deviation on the forward rate is shown in Figure 4.5, which shows the two forward rate curves from Figure 4.4, with two additional forward rate curves where the standard deviation has been raised from 0.02 to 0.05.

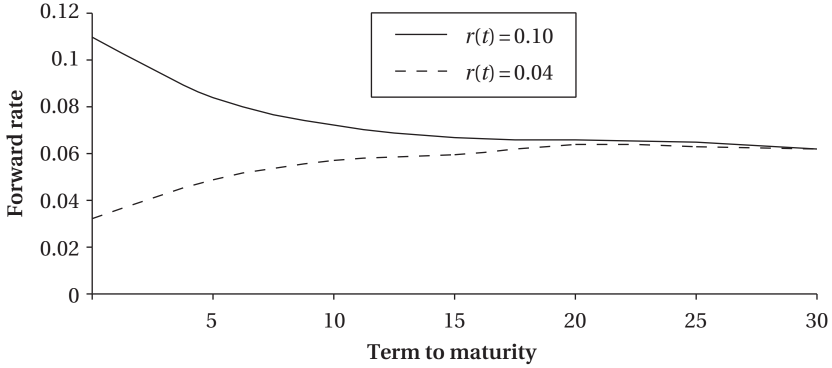


Figure 4.4: Forward rate curves with spot rate $r(t) = 0.04$ and $r(t) = 0.10$.

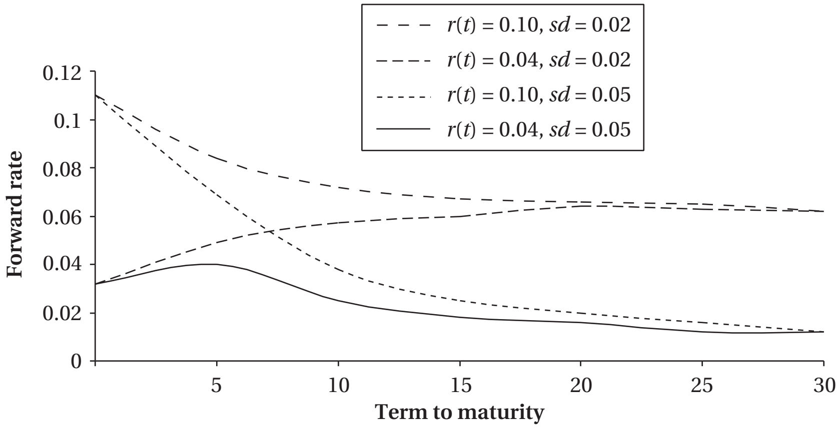


Figure 4.5: Forward rate curves with standard deviations of 0.02 and 0.05.

In describing the dynamics of the yield curve the Vasicek model only captures changes in the short-rate *r*, and not the long-run average rate *b*.

A key point about this model is that as the short-rate follows a normal distribution, it has a positive probability of becoming negative at any point in time. This is common to all models that assume a Gaussian interest-rate process, and although it might be considered a significant drawback, in fact it will only be exhibited under extreme parameter values. For instance in the example in Figure 4.4 the forward rates are not unusual; however if we increase the standard deviation the effect will be to decrease forward rates, and this ultimately produces negative forward rates. For example if we calculate the forward rates for a standard deviation $\sigma = 0.09$, the result will be to produce negative rates, as shown in Figure 4.6. A negative forward rate is equivalent to a zero-coupon bond price that increases over time, which is clearly unrealistic under all but the most unusual and rare conditions. The reason that Gaussian interest-rate models can produce negative forward rates when the standard deviation is high is because the probability of achieving negative interest rates is high. Under certain parameter values, particularly under high values for the standard deviation, the probability of negative forward rates exists. However we saw that this is only under certain parameters, and in fact the presence of mean reversion makes this a low possibility.

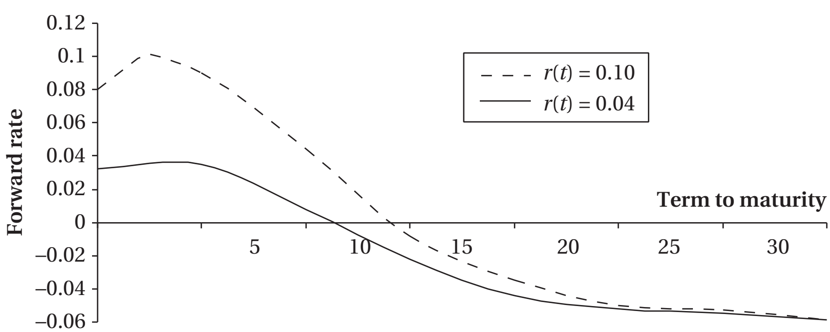


Figure 4.6: Forward rate curves under high volatility.

It might be considered to be more realistic to consider that there are no constant parameters for the drift rate and the standard deviation that would ensure that the price of a zero-coupon bond at any time is exactly the same as that suggested by observed market yields. For this reason a modified version of the Vasicek model has been described by Hull and White (1990), known as the Hull–White or extended Vasicek model, which we consider later.

4.3.2 *The Merton model*

In Merton's model (1971) the interest-rate process is assumed to be a generalised Weiner process, described by (4.31):

$$r(t) = r_0 + \alpha t + \sigma W(t) \tag{4.31}$$

which in differential form is given by (4.32):

$$\mathrm{d}r = \alpha \mathrm{d}t + \sigma \mathrm{d}W. \tag{4.32}$$

For $0 \le t \le T$ it can be shown that

$$r(T) = r(t) + \alpha(T - t) + \sigma(W(T) - W(t)). \tag{4.33}$$

The distribution of $r(T)$ is normal with a mean of $r(t) + \alpha(T - t)$ and standard deviation of $\sigma\sqrt{T - t}$.

For a fixed term to maturity T the forward rate $f(r(t), T - t)$ is an Itô process of the form:

$$\begin{aligned} \mathrm{d}fr(r(t),\ T - t) &= \mathrm{d}r - \alpha \mathrm{d}t + \sigma^2 (T - t)\mathrm{d}t \\ &= \sigma^2 (T - t)\mathrm{d}s + \sigma \mathrm{d}W. \end{aligned} \tag{4.34}$$

The continuously compounded yield at time t of a risk-free zero-coupon bond paying 1 on maturity at time T is given by:

$$\begin{aligned} R(t,\ T) &= \frac{1}{T - t} ln\left(\frac{1}{P(r(t),\ T - t)}\right) \\ &= \frac{1}{T - t} A(T - t) + R(t) \\ &= R(r(t),\ T - t) \end{aligned} \tag{4.35}$$

where R is the function

$$R(r,\ T) = \frac{1}{T} A(T) + r = \frac{1}{2}\alpha T - \frac{\sigma^2}{6} T^2 + r. \tag{4.36}$$

The average future interest rate over the time period $(t,\ T)$ is given by (4.37):

$$r_a = \frac{1}{T - t} \int_t^T r(s)\mathrm{d}s. \tag{4.37}$$

In the Merton model forward rates will always be negative at long maturities, unlike the Vasicek model where there are a range of parameters under which the forward rates will be positive at all maturities. This is because although in both models the forward rate is negatively affected by the standard deviation of the future interest rate, which is an increasing function of the time to maturity, in the Merton model it changes in a linear fashion to infinity, whereas in the Vasicek model it grows to a finite limit. Therefore the standard deviation is more powerful in the Merton model, and it results in the forward rates being negative at long maturities.

4.3.3 *The Cox–Ingersoll–Ross model*

From the previous section we see that under a model that assumes the short-rate to follow a normal distribution, there can arise instances of negative forward rates. The Cox–Ingersoll–Ross model (1985) is a one-factor model and as originally presented removed the possibility of negative rates.[8] Under the CIR model the dynamics of the short-rate are described by (4.38):

$$\mathrm{d}r = a(b - r)\mathrm{d}t + \sigma\sqrt{r}\mathrm{d}W \tag{4.38}$$

which like Vasicek also captures a mean-reverting phenomenon. However the stochastic term has a standard deviation that is proportional to $\sqrt{r}$. This is a significant difference because it states that as the short-rate increases, the standard deviation will decrease. This means that forward rates will be positive. In the CIR model the price of a risk-free zero-coupon bond is given by:

$$P(t,\ T) = A(t,\ T)e^{-B(t,\ T)r} \tag{4.39}$$

where

$$\begin{aligned} B(t,\ T) &= \frac{2(e^{\gamma(T-t)} - 1)}{(\gamma + a)(e^{\gamma(T-t)} - 1) + 2\gamma} \\ A(t,\ T) &= \left(\frac{2\gamma e^{(a+\gamma)(T-t)/2}}{(\gamma + a)(e^{\lambda(T-t)} - 1) + 2\gamma}\right)^{2ab/\sigma^2} \\ \gamma &= \sqrt{a^2 + 2\sigma^2}. \end{aligned}$$

The long-run interest rate $R(t, T)$ is a function of the short-rate $r(t)$, so that the short-rate only is all that is required to fit the entire term structure.

4.3.4 *General comment*

The Gaussian models, also called affine models (see for example, James and Webber, 2000) are popular because they are straightforward to implement and they provide explicit numerical solutions when used in instrument pricing. Although Gaussian models allow negative interest rates under certain conditions, this is not necessarily a completely unrealistic trait, although some academic opinion holds that any model that allows negative interest rates cannot be correct and should not be used. Negative interest rates will only result under very specific conditions, which have a low probability (see for example Rogers, 1995), and for this reason these models remain popular. However in an environment of low interest rates for instance, the CIR type models, which do not permit negative interest rates, may be preferred.

4.4 Arbitrage-free models

An equilibrium model of the term structure, of which we reviewed three in the previous section, is a model that is derived from (or consistent with) a general equilibrium model of the economy. They use generally constant parameters, including most crucially a constant volatility, and the actual parameters used are often calculated from historical time series

8 Although formally published in 1985, the Cox–Ingersoll–Ross model was being circulated in academic circles from the mid-1970s onwards, which would make it one of the earliest interest-rate models.

data. Banks commonly also use parameters that are calculated from actual data and implied volatilities, which are obtained from the prices of exchange-traded option contracts.

An arbitrage-free model of the term structure on the other hand, can be made to fit precisely with the current, observed term structure, so that observed bond yields are in fact equal to the bond yields calculated by the model. So an arbitrage-free model is intended to be consistent with the currently observed zero-coupon yield curve, and the short-rate drift rate is dependent on time, because the future average path taken by the short-rate is determined by the shape of the initial yield curve. This means that in a positively sloping yield curve environment, the short-rate r will be increasing, on average, while it will be decreasing in an initial inverted yield curve environment. In a humped initial yield curve environment, the expected short-rate path will also be humped. In an equilibrium model the drift term for the short rate (that is, the coefficient of the $\mathrm{d}t$ term given above) is not dependent on time.

In theory, the price predicted by any model, were it to be observed in the market, would render that model to be an *arbitrage-free* one; however arbitrage-free models are so-called because they compare the model-predicted price to the actual market price. In an equilibrium model the initial term structure is a product of the model, while in an arbitrage-free model the actual term structure is an input to the model in the first place. In practice an equilibrium model may not be arbitrage-free under certain conditions; namely it may show small errors at particular points along the curve, or it may feature a large error across the whole term structure. The most fundamental issue in this regard is that the concept of the risk-free short-term interest rate is difficult to identify as an actual interest rate in the money market. In practice there may be more than one interest rate that presents itself, for example the T-bill rate or the same maturity government bond repo rate, and this remains a current issue.

For these reasons practitioners may prefer to use an arbitrage-free model if one can be successfully implemented and calibrated. This is not always straightforward, and under certain conditions it is easier to implement an equilibrium multi-factor model (which we discuss in the next section) than it is to implement a multi-factor arbitrage-free model. Under one particular set of circumstances however it is always preferable to use an equilibrium model, and that is when reliable market data is not available. If modelling the term structure in a developing or 'emerging' bond market, it will be more efficient to use an equilibrium model.

Some texts have suggested that equilibrium models can be converted into arbitrage-free models by making the short-rate drift rate time dependent. However this may change the whole nature of the model, presenting problems in calibration.

4.4.1 *The Ho and Lee model*

The Ho–Lee model (1986) was one of the first arbitrage-free models and was presented using a binomial lattice approach, with two parameters; the standard deviation of the short-rate, and the risk premium of the short-rate. We summarise it here. Following Ho and Lee, let $P_i^{(n)}(\cdot)$ be the equilibrium price of a zero-coupon bond maturing at time T under state i. That is $P_i^{(n)}(\cdot)$ is a discount function describing the entire term structure of interest rates, and will satisfy the following conditions:

$$P_i^{(n)}(0) = 1$$
$$\lim_{T\to\infty} P_i^{(n)}(T) = 0.$$

To describe the binomial lattice we denote the price at the initial time 0 as $P_0^{(0)}(\cdot) = 1$.

At time 1 the discount function is specified by two possible functions $P_1^{(1)}(0)$ and $P_0^{(1)}(0)$ which correspond respectively to the upside and the downside outcomes. Therefore at time n the binomial process is given by the discount function $P_i^{(n)}(\cdot)$ which can move upwards to a function $P_{i+1}^{(n+1)}(\cdot)$ and downwards to a function $P_i^{(n+1)}(\cdot)$ for $i=0$ to n.

As described by Ho and Lee there are two functions denoted $h(T)$ and $h^*(T)$ that describe the upstate and downstate as (4.40) and (4.41) respectively, below,

$$P_{i+1}^{(n+1)}(T) = \left(\frac{P_i^{(n)}(T+1)}{P_i^{(n)}(1)}\right)h(T) \tag{4.40}$$

$$P_{i+1}^{(n+1)}(T) = \left(\frac{P_i^{(n)}(T+1)}{P_i^{(n)}(1)}\right)h^*(T) \tag{4.41}$$

with $h(0)=h^*(0)=1$.

The two functions specify the deviations of the discount functions from the implied forward functions. To satisfy arbitrage-free conditions, they define an implied binomial probability π that is independent of time T, while the initial discount function $P(T)$ is given by:

$$\pi h(T) + (1-\pi)h^*(T) = 1 \quad \text{for} \quad n, i > 0 \tag{4.42}$$

and

$$P_i^{(n)}(T) = \left(\pi P_{i+1}^{(n+1)}(T-1) + (1-\pi)P_i^{(n+1)}(T-1)\right)P_i^{(n)}(1). \tag{4.43}$$

Equation (4.43) shows that the bond price is equal to the expected value of the bond, discounted at the prevailing one-period rate. Therefore π is the implied risk-neutral probability.

The assumption that the discount function evolves from one state to another as a function only of the number of upward and downward movements is equivalent to the assumption that a downward movement followed by an upward movement is equivalent to an upward movement followed by a downward movement. This produces the values for h and h^* given by (4.44).

$$h(T) = \frac{1}{\pi + (1-\pi)\delta^T} \quad \text{for} \quad T \geq 0 \tag{4.44}$$

$$h^*(T) = \frac{\delta^T}{\pi + (1-\pi)\delta^T} \tag{4.45}$$

where δ is the interest-rate spread.

It has been shown[9] that the model describes a continuous time process given by

$$\mathrm{d}r = \theta(t)\mathrm{d}t + \sigma \mathrm{d}W(t) \tag{4.46}$$

where σ is the constant instantaneous standard deviation of the short-rate and $\theta(t)$ is a time-dependent function that describes the short-rate process and fits the model to the current observed term structure. This term defines the average direction that the short-rate moves at time t, which is independent of the short-rate. The variable $\theta(t)$ is given by:

$$\theta(t) = f(0,\ t) + \sigma^2 t \tag{4.47}$$

[9] For example, see Hull (1997).

where $f(0, t)$ is the instantaneous forward rate for the period t at time 0. In fact the term $\theta(t)$ approximates to $f(0, t)$ which states that the average direction of the short-rate in the future is given by the slope of the instantaneous forward curve.

The Ho and Lee model is straightforward to implement and is regarded by practitioners as convenient because it uses the information available from the current term structure, so that it produces a model that precisely fits the current term structure. It also requires only two parameters. However it assigns the same volatility to all spot and forward rates, so the volatility structure is restrictive for some market participants. In addition the model does not incorporate mean reversion.

4.4.2 *The Hull–White model*

The Hull–White model (1990) is an extension of the Vasicek model designed to produce a precise fit with the current term structure of rates. It is also known as the *extended Vasicek model*, with the interest rate following a process described by (4.48):

$$\mathrm{d}r = (\alpha - ar)\mathrm{d}t + \sigma \mathrm{d}W(t). \tag{4.48}$$

It is also sometimes written as

$$\mathrm{d}r = a\left(\frac{\alpha}{a} - r\right)\mathrm{d}t + \sigma \mathrm{d}W(t) \tag{4.49}$$

where a is the mean reversion rate and a and σ are constants. It has been described as a Vasicek model with a time-dependent reversion level. The model is also called the *general Hull–White model*, while a special case where $a \neq 0$ is known as the simplified Hull–White model. In the Vasicek model $a \neq 0$ and $\alpha = ab$ where b is constant.

The Hull–White model can be fitted to an initial term structure, and also a volatility term structure. A comprehensive analysis is given in Pelsser (1996) as well as James and Webber (2000).

It can be shown that

$$r(t) = e^{-K(t)}\left(r_0 + \int_t^T e^K \alpha \mathrm{d}t + \int_t^T e^K \sigma \mathrm{d}W(t)\right) \tag{4.50}$$

where the process K is given by $K(t) = \int_t^T a\mathrm{d}t$.

To calculate the price of a zero-coupon bond, the first step is to calculate the integral $I(t, T) = \int_t^T r\mathrm{d}s$ which follows a normal distribution with mean $m(r(t), t; T)$ and standard deviation $\sqrt{\nu(t; T)}$. The price of a bond is given by (4.51):

$$\begin{aligned} P(t,\ T) &= E_Q[exp[-I(t,\ T)]|F_t] \\ &= exp\left(-m(r(t),\ t; T) + \frac{1}{2}\nu(t; T)\right) \\ &= P(r(t),\ t; T) \end{aligned} \tag{4.51}$$

where $P(r, t; \mathrm{T})$ is the function

$$P(r(t), t; T) = exp\left(-m(r, t; T) + \frac{1}{2}\nu(t; T)\right). \tag{4.52}$$

The price of a zero-coupon discount at time t is defined in terms of the short-rate r at time t and the current term structure. The price function is not static, and the price of a bond at time t that matures at time T is a function of the short-rate, as we have noted, and separately of the time t.

The volatility of the bond price is given by the function $B(t; T)\sigma(t)$ where B is defined as

$$B(t,\ T) = \int_t^T e^{K(u)+K(t)}\mathrm{d}u = e^{K(t)}\int_t^T e^{-K}\mathrm{d}u. \tag{4.53}$$

The bond price volatility is a deterministic function of t. The 'pull to par' of the zero-coupon bond is captured by the fact that the volatility reduces to zero as t approaches T, as long as σ is continuous at t. As the mean m is normally distributed, it follows that the bond price is lognormally distributed, so therefore we have the function

$$ln\ P(t,\ T) = -A(t,\ T) - B(t,\ T)r(t)$$

where $A(t, T)$ is defined by $A(t,\ T) = \int_t^T e^{-K(u)} \int_t^u e^K \alpha \mathrm{d}x\mathrm{d}u - \frac{1}{2}\nu(t,\ T)$.

The price function above can be continuously differentiated as a function of t. The forward rate is given by (4.54):

$$\begin{aligned} f(t,\ T) &= -\frac{\partial}{\partial T} ln\ P(t,\ T) \\ &= A_T(t,\ T) + B_T(t,\ T)r(t) \\ &= e^{-K(T)}\int_t^T \alpha \mathrm{d}x - \frac{1}{2}\nu_T(t,\ T) + e^{-K(T)+K(t)}r(t) \\ &= f(r(t); t,\ T) \end{aligned} \tag{4.54}$$

where $f(r(t); t, T)$ is defined by the function below:

$$f(r(t); t,\ T) = e^{-K(T)}\int_t^T \alpha \mathrm{d}x - \frac{1}{2}\nu_T(t,\ T) + e^{-K(T)+K(t)}r. \tag{4.55}$$

The forward rate function f at time t is not static and is a function of the short-rate r at time t, the time t and the time to maturity T. The Hull–White model can be calibrated in terms of the forward rate f. That is, at time t the information (parameters) required to implement this are the short-rate $r(t)$, the standard deviation σ of the short-rate, the forward rate f and the standard deviations $B_T(t,\ T)\sigma(t)$ of the forward rates at time t. If the forward rates are known in a form that allows their first differential to be calculated with respect to t, using the other information it is possible to calculate the function B_T, the derivative of this function and thereby the value for $a(t)$, using the relationship in (4.56):

$$a(t) = -\frac{B'_T(t,\ T)}{B_T(t,\ T)} \tag{4.56}$$

which describes the volatility of the bond price as a function of the maturity date T.

The continuously compounded yield of a zero-coupon bond at time t that matures at time T is shown to be

$$\begin{aligned} R(t,\ T) &= \frac{1}{T-t}\left(m(r(t), t; T) - \frac{1}{2}\nu(t,\ T)\right) \\ &= \frac{1}{T-t} ln\left(\frac{1}{P(t,\ T)}\right) \end{aligned} \tag{4.57}$$

$$= \frac{1}{T-t}(A(t,\ T) + B(t,\ T)r(t))$$
$$= R(r(t),\ t,\ T)$$

where R is given by the function shown in (4.58):

$$R(r,\ t,\ T) = \frac{1}{T-t}(A(t,\ T) + B(t,\ T)r). \tag{4.58}$$

Like the bond price function, the yield on a zero-coupon bond is a function of the short-rate r and follows a normal distribution; the yield curve is a function of the short-rate r, the time t and the time to maturity T. The long-run average future interest over the time to maturity $(t,\ T)$ is normally distributed and given by:

$$r_a = \frac{1}{T-t}\int_t^T r(s)\mathrm{d}s = \frac{1}{T-t}I(t,\ T). \tag{4.59}$$

4.4.3 *The Black–Derman–Toy model*

In the models we have reviewed in this chapter there has only been one function of time, the parameter α. In certain models either or both of the parameters a and σ are also made to be functions of time. In their 1990 paper Black, Derman and Toy (BDT) proposed a binomial lattice model described by (4.60),

$$\mathrm{d}\ ln\ r = \left(\alpha + \frac{\sigma'(t)}{\sigma(t)}ln(r)\right)\mathrm{d}t + \sigma(t)\mathrm{d}W \tag{4.60}$$

where $\sigma'(t)$ is the partial derivative of σ with respect to time t. The BDT model is a lognormal model, which means that the short-rate volatility is proportional to the instantaneous short-rate, so that the ratio of the volatility to the rate level is constant. The drift term is more complex than that described in the earlier models, and so the BDT model requires numerical fitting to the observed current interest-rate and volatility term structures. That is, the drift term is not calculated analytically. The short-rate volatility is also linked to the mean reversion such that where long-term rates are less volatile than the short-rate, the short-rate volatility will decrease in the long-term. A later model developed by Black and Karasinski (1991) removed the relationship between mean reversion and the volatility level. This is given in (4.61):

$$\mathrm{d}\ ln\ r = (\theta(t) - a(t)ln(r))\mathrm{d}t + \sigma(t)\mathrm{d}W(t). \tag{4.61}$$

As with the previous models the key factor is the short-rate. Using the binomial tree approach, a one-step tree is used to derive the current short-rate to the short-rates one period in the future. These derived rates are then used to derive rates two periods away, and so on.

4.5 Fitting the model

Implementing an interest-rate model requires the input of the term structure yields and volatility parameters, which are used in the process of calibrating the model. The process of fitting the model is called *calibration*. This can be done in at least three ways, which are:

- calibration to the current spot rate yield curve, using a pre-specified volatility level and not the volatility values given by the prices of exchange-traded options. This may result in mispriced bonds and options if the selected volatilities are not accurate;

- calibration to the current spot rate curve and using the volatilities implied by the prices of exchange-traded options; therefore the model would be implemented using volatility parameters that were exactly similar to those implied by the traded option prices. In practice this can be a lengthy process;
- calibrating the model to the current spot rate curve, using volatility parameters that are approximately close enough to result in prices that are near to those of observed exchange-traded options. This is usually the method that is adopted.

Generally volatility values for the different period interest rates are taken from the volatilities of exchange-traded options. However where great accuracy is not required, for example for regulatory capital purposes practitioners may use the first method, while for the purposes of fixed income research the third method is suitable. In both the second and third methods there is the danger that calibrating the model to option prices will result in error simply because the options are mispriced. This is quite possible if using long-dated and/or OTC options, which frequently differ in price according to which bank is pricing them.

In any case a model will usually therefore use volatility inputs from option prices for a range of options that range in maturity from the shortest period to the longest in the term structure. To test the accuracy of the model, one can use the expression in (4.62):

$$\sum_{n=1}^{N}(p_n - P_n)^2 \tag{4.62}$$

where p_n is the observed price of the nth option and P_n is the price of the option as calculated by the model, and N options have been used to calibrate the model. A model that has the lowest value given by (4.62) can be considered to be the most accurate. In deciding which option products should be used to calibrate the model, care should be taken to use instruments that are most similar to the instrument that is being priced by the model.

The different models can lend themselves to a particular calibration method. In the Ho–Lee model, only parallel yield curve shifts are captured and the current yield curve is a direct input; therefore a constant volatility parameter is used. This implies that all the forward rate implied volatilities are identical. In practice this is not necessarily realistic, as long-dated bond prices often experience lower volatility than short-dated bond prices. The model also assumes that volatility is a decreasing function of the time to maturity, which may also be unrealistic. Models that incorporate mean reversion can be implemented with more realistic volatility parameters, as it is the mean reversion effect that results in long-dated bonds having lower volatilities. Therefore a mean-reverting model can be implemented more accurately using the second or third methods described above.

To recap on the issues involved in fitting the extended Vasicek model or Hull–White model; this describes the short-rate process as following the form

$$\mathrm{d}r = (\alpha - ar)\mathrm{d}t + \sigma \mathrm{d}W(t). \tag{4.63}$$

In implementing this model, there are three possible approaches. The model could be calibrated by keeping α and a constant and calibrating the standard deviation parameter. This means that the model is fitted to the current yield curve and the volatility value is adjusted to that required to produce the observed curve. However this may result in high volatility values, which rise by a squared function, and therefore will not be realistic. The second method is to calibrate α, keeping the other two parameters constant. This is adjusting the mean reversion rate in order to fit the derived curve with the observed curve. The resulting derived yield curve will be a function of the current short-rate and the mean reversion rate. This method is sometimes applied in practice, although it can result in inaccurate volatility levels for long-dated bonds, because large adjustments in the mean reversion rate are needed to fit the derived curve to the long-dated part of the observed curve. The third approach would be to calibrate *ar*, keeping the other parameters constant. This produces a stable yield curve and is most commonly followed by practitioners in the market.

4.6 Summary

In this chapter we have considered both equilibrium and arbitrage-free interest-rate models. These are one-factor Gaussian models of the term structure of interest rates. We saw that in order to specify a term structure model, the respective authors described the dynamics of the price process, and that this was then used to price a zero-coupon bond. The short-rate that is modelled is assumed to be a risk-free interest rate, and once this is modelled we can derive the forward rate and the yield of a zero-coupon bond, as well as its price. So it is possible to model the entire forward rate curve as a function of the current short-rate only, in the Vasicek and Cox–Ingersoll–Ross models, among others. Both the Vasicek and Merton models assume constant parameters, and because of equal probabilities of forward rates and the assumption of a normal distribution, they can, under certain conditions relating to the level of the standard deviation, produce negative forward rates.

The models are based on the fact that the price of a bond, which exhibits a pull-to-par effect, and the forward rate, are both Itô processes. For the bond price the relative drift is the interest rate, and is deterministic, as is the forward rate. The bond price, yield and forward rate are functions of the current short-rate, and follow a normal distribution. An increase in the short-rate will result in a rise in the forward rates, and this is more pronounced for the shortest maturity rates. The instantaneous volatility of the forward rates decreases with decreasing time to maturity, and approaches the volatility of the current short-rate at time t.

The Vasicek, Cox–Ingersoll–Ross, Hull–White and other models incorporate mean reversion. As the time to maturity increases and as it approaches infinity, the forward rates converge to a point at the long-run mean reversion level of the current short-rate. This is the limiting level of the forward rate and is a function of the volatility of the current short-rate. As the time to maturity approaches zero, the short-term forward rate converges to the same level as the instantaneous short-rate. In the Merton and Vasicek models the mean of the short-rate over the maturity period T is assumed to be constant. The same constant for the mean, or the *drift* of the interest rate, is described in the Ho–Lee model, but not the extended Vasicek or Hull–White model.

We also noted that the efficacy of a model was not necessarily solely related to how realistic its assumptions might be, but how straightforward it was to implement in practice, that is, the ease with which it could be calibrated.

Appendix

Appendix 4.1: Illustration of forward rate structure when spot rate structure is increasing

We assume the spot rate $r(0, T)$ is a function of time and is increasing to a high point at $\overline{T}$. It is given by

$$r(0,\ T) = \frac{\int_0^T f(0,\ s)\mathrm{d}s}{T}. \tag{4.64}$$

At its high point the function is neither increasing nor decreasing, so we may write

$$\frac{\mathrm{d}r(0,\ \overline{T})}{\mathrm{d}T} = 0 \tag{4.65}$$

and therefore the second derivative with respect to T will be

$$\frac{\mathrm{d}^2 r(0,\ \overline{T})}{\mathrm{d}T^2} < 0 \tag{4.66}$$

From (4.14) and (4.65) we may state

$$f(0,\ \overline{T}) = r(0,\ \overline{T}) \tag{4.67}$$

and from (4.66) and (4.67) the second derivative of the spot rate is

$$\frac{\mathrm{d}^2 r(0,\ \overline{T})}{\mathrm{d}T^2} = \left[\frac{\mathrm{d}f(0,\ \overline{T})}{\mathrm{d}T} - \frac{\mathrm{d}r(0,\ \overline{T})}{\mathrm{d}T}\right]\frac{1}{\overline{\overline{T}}} < 0. \tag{4.68}$$

From (4.65) we know the spot rate function is zero at $\overline{T}$ so the derivative of the forward rate with respect to T would therefore be

$$\frac{\mathrm{d}f(0,\ \overline{T})}{\mathrm{d}T} < 0. \tag{4.69}$$

So in this case the forward rate is decreasing at the point $\overline{T}$ when the spot rate is at its maximum value. This is illustrated hypothetically in Figure 4.1 and it is common to observe the forward rate curve decreasing as the spot rate is increasing.

Selected bibliography and references

Baxter, M., Rennie, A., *Financial Calculus*, Cambridge University Press, 1996.

Black, F., Derman, E., Toy, W., 'A one-factor model of interest rates and its application to Treasury bond options', *Financial Analysts Journal*, Jan–Feb 1990, pp. 33–39.

Black, F., Karasinski, P., 'Bond and option prices when short rates are lognormal', *Financial Analysts Journal*, Jul–Aug 1991, pp. 52–59.

Brennan, M., Schwartz, E., 'A continuous time approach to the pricing of bonds', *Journal of Banking and Finance* 3, 1979, p. 134 *ff*.

Brennan, M., Schwartz, E., 'Conditional predictions of bond prices and returns', *Journal of Finance* 35, 1980, p. 405 *ff*.

Campbell, J., Lo, A., MacKinlay, A., *The Econometrics of Financial Markets*, Princeton University Press, 1997.

Chan, K., *et al.*, 'An empirical comparison of alternative models of the short-term interest rate', *Journal of Finance* 47, 1992, pp. 1209–1227.

Choudhry, M., *Bond Market Securities*, FT Prentice Hall, 2001.

Cox, J., Ingersoll, J., Ross, S., 'A theory of the term structure of interest rates', *Econometrica* 53, 1985, pp. 385–407.

Duffie, D., *Dynamic Asset Pricing Theory*, 2nd edition, Princeton University Press, 1996.

Fisher, L., Leibowitz, M., 'Effects of alternative anticipations of yield curve behaviour on the composition of immunized portfolios and on their target returns', in Kaufmann, G., *et al.* (eds.), *Innovations in Bond Portfolio Management*, Jai Press, 1983.

Ho, T., Lee, S., 'Term structure movements and pricing interest rate contingent claims', *Journal of Finance* 41, 1986, pp. 1011–1029.

Hull, J., *Options, Futures and Other Derivatives*, Wiley, 1997.

Hull, J., White, A., 'Pricing interest-rate derivative securities', *Review of Financial Studies* 3, 1990, pp. 573–592.

James, J., Webber, N., *Interest Rate Modelling*, Wiley, 2000.

Jamshidian, F., 'Bond, futures and option valuation in the quadratic interest rate model', *Applied Mathematical Finance* 3, 1996, pp. 93–115.

Jarrow, R., *Modelling Fixed Income Securities and Interest Rate Options*, McGraw-Hill, 1996.

Merton, R., 'Optimum consumption and portfolio rules in a continuous time model', *Journal of Economic Theory* 3, 1971, pp. 373–413.

Merton, R., 'Theory of option pricing', *Bell Journal of Economics and Management Science* 4, 1973, pp. 141–183.

Neftci, S., *An Introduction to the Mathematics of Financial Derivatives*, 2nd edition, Academic Press, 2000.

Pelsser, A., 'Efficient methods for valuing and managing interest rate and other derivative securities', PhD thesis, Erasmus University, Rotterdam, 1996.

Phoa, W., *Advanced Fixed Income Analytics*, FJF Associates, 1998.

Pliska, S., 'A stochastic calculus model of continuous trading: optimal portfolios', *Mathematics of Operations Research* 11, 1986, pp. 371–382.

Rebonato, R., *Interest Rate Option Models*, 2nd edition, Wiley, 1998.

Rogers, L., 'Which model for the term-structure of interest rates should one use?', in Davis, M., *et al.* (eds.), *Mathematical Finance*, Springer-Verlag, 1995, pp. 93–115.

Steeley, J., 'Estimating the gilt-edged term structure: basis splines and confidence intervals', *Journal of Business Finance and Accounting* 18, 1991, pp. 513–530.

The Economist, 'Out of Debt', *The Economist*, 12 February 2000.

Van Deventer, J., Imai, K., *Financial Risk Analysis*, McGraw-Hill, 1997.

Vasicek, O., 'An equilibrium characterization of the term structure', *Journal of Financial Economics* 5, 1977, pp. 177–188.

5 Interest-rate Models II

In this chapter we consider multi-factor and whole yield curve models. As we noted in the previous chapter, short-rate models have certain drawbacks, which though not necessarily limiting their usefulness, do leave room for further development. The drawbacks are that, as the single short-rate is used to derive the complete term structure, in practice this can be unsuitable for the calculation of bond yields. When this happens it becomes difficult to visualise the actual dynamics of the yield curve, and the model no longer fits observed changes in the curve. This means that the accuracy of the model cannot be observed. Another drawback is that in certain equilibrium model cases the model cannot be fitted precisely to the observed yield curve, as they have constant parameters. In these cases calibration of the model is on a 'goodness of fit' or best fit approach.

In response to these issues interest-rate models have been developed that model the entire yield curve. In a whole yield curve, the dynamics of the entire term structure are modelled. The Ho–Lee model is a simple type of whole curve model, which allows random parallel shifts in the yield curve. More advanced models, the Heath–Jarrow–Morton family of models, are discussed in this chapter, as are factors involved in their implementation.

5.1 Introduction

A landmark development in interest-rate modelling has been the specification of the dynamics of the complete term structure. In this case the volatility of the term structure is given by a specified function, which may be a function of time, term to maturity or zero-coupon rates. A simple approach is described in the Ho–Lee model, in which the volatility of the term structure is a parallel shift in the yield curve, the extent of which is independent of the current time and the level of current interest rates. The Ho–Lee model is not widely used, although it was the basis for the Heath–Jarrow–Morton (HJM) model, which *is* widely used. The HJM model describes a process whereby the whole yield curve evolves simultaneously, in accordance with a set of volatility term structures. The model is usually described as being one that describes the evolution of the forward rate, however it can also be expressed in terms of the spot rate or of bond prices (see for example James and Webber (2000), Chapter 8). For a more detailed description of the HJM framework refer to Baxter and Rennie (1996), Hull (1997), Rebonato (1998), Bjork (1996) and James and Webber (2000). Baxter and Rennie is very accessible, while Neftci (1996) is an excellent introduction to the mathematical background.

In seeking to develop a model for the entire term structure, the requirement is to model the behaviour of the entire forward yield curve, that is the behaviour of the forward short-rate $f(t, T)$ for all forward dates T. Therefore we require the random process $f(T)$ for all forward dates T.

Given this, it can be shown that the yield R on a T-maturity zero-coupon bond at time t is the average of the forward rates at that time on all the forward dates s between t and T, given by (5.1):

$$R(t,\ T) = \frac{1}{T-t}\int_t^T f_t(s)\mathrm{d}s. \tag{5.1}$$

To model the complete curve it is necessary to specify a drift rate and volatility level for $f(t,\ T)$ for each T.

5.2 The Heath–Jarrow–Morton model

A landmark development in the longstanding research into yield curve modelling was presented by David Heath, Robert Jarrow and Andrew Morton in their 1989 paper, which formally appeared in volume 60 of *Econometrica* (1992). The paper considered interest-rate modelling as a stochastic process, but applied to the entire term structure rather than only the short-rate. The importance of the Heath–Jarrow–Morton (HJM) presentation is this: in a market that permits no arbitrage, where interest rates including forward rates are assumed to follow a Weiner process, the drift term and the volatility term in the model's stochastic differential equation are not independent from each other, and in fact the drift term is a deterministic function of the volatility term. This has significant practical implications for the pricing and hedging of interest-rate options.

The general form of the HJM model is very complex, not surprisingly as it is a multi-factor model. We begin by describing the single-factor HJM model. This section is based on Chapter 5 of Baxter and Rennie, *Financial Calculus*, Cambridge University Press (1996), and follows their approach with permission. This work is an accessible and excellent text and is highly recommended.

5.2.1 *The single-factor HJM model*

In the previous chapter, and indeed in previous analysis, we have defined the *forward rate* as the interest rate applicable to a loan made at a future point in time and repayable instantaneously. We assume that the dynamics of the forward rate follow a Weiner process. The *spot rate* is the rate for borrowing undertaken now and maturing at T, and we know from previous analysis that it is the geometric average of the forward rates from 0 to T that is

$$r(0,\ T) = T^{-1}\int_t^T f(0,\ t)\mathrm{d}t. \tag{5.2}$$

We also specify a money market account that accumulates interest at the continuously compounded spot rate r.

A default-free zero-coupon bond can be defined in terms of its current value under an *initial probability measure*, which is the Weiner process that describes the forward rate dynamics, and its price or present value under this probability measure. This leads us to the HJM model, in that we are required to determine what is termed a 'change in probability measure', such that the dynamics of the zero-coupon bond price are transformed into a *martingale*. This is carried out using Itô's lemma and a transformation of the differential equation of the bond price process. It can then be shown that in order to prevent arbitrage there would have to be a relationship between drift rate of the forward rate and its volatility coefficient.

First we look at the forward rate process. We know from the previous chapter for $[0, T]$ at time t that the stochastic evolution of the forward rate $f(t, T)$ can be described as

$$\mathrm{d}f(t,\ T) = a(t,\ T)\mathrm{d}t + \sigma(t,\ T)\mathrm{d}W_t \tag{5.3}$$

or alternatively in integral form as

$$f(t,\ T) = f(0,\ T) + \int_0^t a(s,\ T)\mathrm{d}s + \int_0^t \sigma(s,\ T)\mathrm{d}W_s \tag{5.4}$$

where a is the drift parameter, σ the volatility coefficient and W_t is the Weiner process or Brownian motion. The terms dz or dZ are sometimes used to denote the Weiner process.

In (5.3) the drift and volatility coefficients are functions of time t and T. For all forward rates $f(t, T)$ in the period $[0, T]$ the only source of uncertainty is the Brownian motion. In practice this would mean that all forward rates would be perfectly positively correlated, irrespective of their terms to maturity. However if we introduce the feature that there is more than one source of uncertainty in the evolution of interest rates, it would result in less than perfect correlation of interest rates, which is what is described by the HJM model.

Before we come to that however we wish to describe the spot rate and the money market account processes. In (5.4) under the particular condition of the maturity point T as it tends towards t (that is $T \to t$), the forward rate tends to approach the value of the short rate (spot rate), so we have

$$\lim_{T\to t} f(t,\ T) = f(t,\ t) = r(t)$$

so that it can be shown that

$$r(t) = f(0,\ t) + \int_0^t a(s,\ t)\mathrm{d}s + \int_0^t \sigma(s,\ t)\mathrm{d}W_s. \tag{5.5}$$

The money market account is also described as a Weiner process. We denote by $M(t, t) \equiv M(t)$ the value of the money market account at time t, which has an initial value of 1 at time 0 so that $M(0, 0) = 1$. This account earns interest at the spot rate $r(t)$ which means that at time t the value of the account is given by

$$M(t) = e^{\int_0^t r(s)\mathrm{d}s} \tag{5.6}$$

that is the interest accumulated at the continuously compounded spot rate $r(t)$. It can be shown by substituting (5.5) into (5.6), that

$$M(t) = exp\left(\int_0^t f(0,\ s)\mathrm{d}s + \int_0^t \int_0^s a(u,\ s)\mathrm{d}u\mathrm{d}s + \int_0^t \int_0^s \sigma(u,\ s)\mathrm{d}W_u\mathrm{d}s\right). \tag{5.7}$$

To simplify the description we write the double integrals in (5.7) in the form given below, which is

$$\int_0^t \int_s^t a(s,\ u)\mathrm{d}u\mathrm{d}s + \int_0^t \int_s^t \sigma(s,\ u)\mathrm{d}u\mathrm{d}W_s.$$

For reasons of space the description of the process by which this simplification is achieved is relegated to a page on the author's Web site.

Using the simplification above, it can be shown that the value of the money market account, which is growing by an amount generated by the continuously compounded spot rate $r(t)$, is given by

$$M(t) = exp\left(\int_0^t f(0,\ u)\mathrm{d}u + \int_0^t \int_s^t a(s,\ u)\mathrm{d}u\mathrm{d}s + \int_0^t \int_s^t \sigma(s,\ u)\mathrm{d}u\mathrm{d}W_s\right). \tag{5.8}$$

The expression for the value of the money market account can be used to determine the expression for the zero-coupon bond price, which we denote $P(t,\ T)$. The money market account earns interest at the spot rate $r(t)$, while the bond price is the present value of 1 discounted at this rate. Therefore the inverse of (5.8) is required, which is

$$M^{-1}(t) = e^{-\int_0^t r(u)\mathrm{d}u}. \tag{5.9}$$

Hence the present value at time 0 of the bond $P(t,\ T)$ is

$$P(t,\ T) = e^{-\int_0^t r(u)\mathrm{d}u}P(t,\ T)$$

and it can be shown that as a Weiner process the present value is given by

$$P(t,\ T) = exp\left(-\int_0^T f(0,\ u)\mathrm{d}u - \int_0^t \int_s^T \sigma(s,\ u)\mathrm{d}u\mathrm{d}W_s - \int_0^t \int_s^T a(s,\ u)\mathrm{d}u\mathrm{d}s\right). \tag{5.10}$$

5.2.2 *Transforming the probability measure*

Since the pioneering work of Harrison and Pliska (1981), which recognised that the absence of arbitrage was linked to the existence of a martingale probability measure, the valuation of derivatives has been deemed to require a probability measure that would transform the underlying security process into a martingale. This is the case here, what is required is a change in probability measure such that $P(t,\ T)$ becomes a martingale.

This is done by using Itô's lemma[1] to transform the stochastic differential equation of the price process, and then determine the change in the Brownian differential dW so that there remains no drift term. The first step is to consider the differential of $P(t,\ T)$. We express this in the form

$$P(t,\ T) = e^{-X_t} \tag{5.11}$$

where X_t is a Weiner process described by

$$X_t = \int_0^T f(0,\ u)\mathrm{d}u + \int_0^t \int_s^T \sigma(s,\ u)\mathrm{d}u\mathrm{d}W_s + \int_0^t \int_s^T a(s,\ u)\mathrm{d}u\mathrm{d}t. \tag{5.12}$$

The differential of X_t is written as

$$\begin{aligned} \mathrm{d}X_t &= \int_t^T \sigma(t,\ u)\mathrm{d}u\mathrm{d}W_t + \int_t^T a(t,\ u)\mathrm{d}u\mathrm{d}t \\ &= \nu(t,\ T)\mathrm{d}W_t + \int_t^T a(t,\ u)\mathrm{d}u\mathrm{d}t \end{aligned} \tag{5.13}$$

---

1 See Appendices 3.2 and 3.3.

where $\nu(t,\ T) \equiv \int_t^T \sigma(t,\ u)\mathrm{d}u$ represents the volatility element of the X_t stochastic process. It can be shown by applying Itô's lemma to (5.11) that

$$\mathrm{d}P(t,\ T) = P(t,\ T)\left(-\nu(t,\ T)\mathrm{d}W_t - \int_t^T a(t,\ u)\mathrm{d}u\mathrm{d}t + \frac{\nu^2}{2}(t,\ T)\mathrm{d}t\right). \tag{5.14}$$

To obtain a new probability measure such that $P(t,\ T)$ is transformed into a martingale, following Baxter and Rennie we effect a change in $\mathrm{d}W_t$ such that (5.14) may be expressed as

$$\mathrm{d}P(t,\ T) = -P(t,\ T)\nu(t,\ T)\mathrm{d}\tilde{W}_t. \tag{5.15}$$

It can then be shown that the solution of (5.15) under such conditions is indeed a martingale, described by

$$P(t) = P_0 exp\left(\int_0^t \nu(\tau,\ T)\mathrm{d}\tilde{W}_\tau - \frac{1}{2}\int_0^t \nu^2(\tau,\ T)\mathrm{d}\tau\right). \tag{5.16}$$

To reiterate, then, we require a transformation of (5.14) so that it becomes a Weiner process with no drift term, in other words a relationship between $\mathrm{d}\tilde{W}_t$ and $\mathrm{d}W_t$. It has been shown that this exists in the form

$$\mathrm{d}\tilde{W}_t = \mathrm{d}W_t + \frac{1}{\nu(t,\ T)}\int_t^T a(t,\ u)\mathrm{d}u\mathrm{d}t - \frac{\nu(t,\ T)}{2}\mathrm{d}t \equiv \mathrm{d}W_t + \gamma_t\mathrm{d}t \tag{5.17}$$

where the change of measure γ_t is given by

$$\gamma_t = \frac{1}{\nu(t,\ T)}\int_t^T a(t,\ u)\mathrm{d}u - \frac{\nu(t,\ T)}{2}. \tag{5.18}$$

5.2.3 *The principle of no-arbitrage*

The demonstration of the no-arbitrage condition in the evolution of the HJM model is perhaps its most significant aspect, as it demonstrates that for arbitrage to be avoided, the volatility function must be related to the drift parameter. This is effected through a constraint that is the change of measure element we introduced just now: again, following Baxter and Rennie and to summarise the original paper, in order to prevent arbitrage, a bond of maturity less than T must have the same change of measure γ_t. The change of measure must by implication therefore not be a function of and be independent from T. It can be shown that if we multiply (5.18) by $\nu(t,\ T)$ and then differentiate it with respect to T we obtain (5.19):

$$\begin{aligned} a(t,\ T) &= \frac{\partial\nu(t,\ T)}{\partial T}(\nu(t,\ T) + \gamma_t) \\ &= \sigma(t,\ T)(\nu(t,\ T) + \gamma_t). \end{aligned} \tag{5.19}$$

The expressions (5.18) and (5.19) represent the two fundamental constraints of the single-factor HJM model. (5.18) is the result of the change in the drift term required by the transformation of $P(t,\ T)$ into a martingale, while (5.19) comes from the need to incorporate a no-arbitrage condition. This is the model in essence, an expression for the value of the drift parameter for $f(t,\ T)$ in the context of the W_t Brownian motion. This impacts as follows: in (5.3) the Brownian motion term $\mathrm{d}W_t$ is replaced by $\mathrm{d}\tilde{W}_t - \gamma_t\mathrm{d}t$ and $a(t,\ T)$ is replaced with the constraint given by (5.19). This results in

$$\begin{aligned} \mathrm{d}f(t,\ T) &= \sigma(t,\ T)(\mathrm{d}\tilde{W}_t - \gamma_t \mathrm{d}t) + \sigma(t,\ T)(\nu(t,\ T) + \gamma_t)\mathrm{d}t \\ &= \sigma(t,\ T)\mathrm{d}\tilde{W}_t + \sigma(t,\ T)\nu(t,\ T)\mathrm{d}t. \end{aligned} \tag{5.20}$$

In conclusion, then, in the single-factor HJM model under the martingale measure the coefficient of the drift must be equal to (5.21):

$$\sigma(t,\ T)\nu(t,\ T) = \sigma(t,\ T)\int_t^T \sigma(t,\ u)\mathrm{d}u. \tag{5.21}$$

In an important application of the HJM model Jarrow and Turnbull (1996) express the price of a zero-coupon risk-free bond as a function of the spot rate $r(t)$, given by

$$P(t,\ T) = \frac{P(0,\ T)}{P(0,\ t)} exp\left(-(r(t) - f(0,\ t))X(t,\ T) - \frac{\sigma^2}{4\lambda}X^2(t,\ T)(1 - e^{-2\lambda t})\right) \tag{5.22}$$

where $X(t,\ T) \equiv (1 - e^{T-t})/\lambda$ and σ and λ are positive constants of the volatility coefficient $\sigma(t, T)$ which is of the form $\sigma exp(-\lambda(T-t))$.

Thus at time t the price of any bond of maturity T is given by the ratio of prices of bonds of maturity t and T, which are the first part of (5.22) and which are observed in the market, and the random variable given by the exponential element of (5.22). If we express the latter as $A(t, T)$ we have

$$P(t,\ T) = \frac{P(0,\ T)}{P(0,\ t)} A(t,\ T) \tag{5.23}$$

and under conditions of zero volatility where $\sigma(t,\ T) = 0$ it can be shown that this element disappears and

$$\begin{aligned} P(t,\ T) &= \frac{P(0,\ T)}{P(0,\ t)} = \frac{e^{-rT}}{e^{-rt}} \\ &= e^{-r(T-t)} \end{aligned} \tag{5.24}$$

which is exactly what we expect.

5.3 Multi-factor term structure models

Previously we considered one-factor models used to varying degrees in the market; these describe only a single kind of change in the yield curve, the parallel shift. In practice there are a range of changes that may occur to the curve, including non-parallel (pivotal) shifts and changes in the slope of the curve. Certain two-factor and multi-factor models have been developed that seek to describe the different type of yield curve shifts. An early two-factor model was that presented by Brennan and Schwartz (1982) which described the stochastic process of the short-rate r and the yield of the long-dated government bond R. In the model these two factors move independently of each other, thus permitting both parallel and pivotal changes in the yield curve. The Brennan–Schwartz model is categorised as a *consol* model in James and Webber (2000). A modified version of the Brennan–Schwartz model[2] has been developed in which the two variable factors are the price of the long bond $P = 1/R$ and the spread between the long-dated yield and the short-rate, which is $S = R - r$. Both factors are assumed to follow a random stochastic process described by (5.25):

$$\begin{aligned} dP &= \mu_P P dt + \sigma_P P dW_t^{(1)} \\ dS &= \mu_S dt + \sigma_S W_t^{(2)} \end{aligned} \tag{5.25}$$

where $E[dW_t^{(1)} dW_t^{(2)}] = p dt$, and the values of μ_P and μ_S are set to be arbitrage-free in terms of the price of bonds of different maturities. So the price of the long-dated bond follows a lognormal process, while the spread S follows a Gaussian process. This means that the spread can be either positive or negative, which permits both a positive sloping or an inverted yield curve.

The modified Brennan–Schwartz model is used in the markets and describes a realistic process for changes in the yield curve and is relatively straightforward to implement; only two variables are required to model the entire term structure.

The Heath–Jarrow–Morton model (1992) is a general approach which is a multi-factor whole yield curve model, where arbitrary changes in the entire term structure can be one of the factors. In practice because of the mass of data that is required to derive the yield curve, the HJM model is usually implemented by means of Monte Carlo simulation, and requires powerful computing systems. The model is described in the next section.

5.3.1 *The multi-factor Heath–Jarrow–Morton model*

A multi-factor model of the whole yield curve has been presented by Heath, Jarrow and Morton (1992).[3] This is a seminal work and a ground-breaking piece of research. The approach models the forward curve as a process arising from the entire initial yield curve, rather than the short-rate only. The spot rate is a stochastic process and the derived yield curve is a function of a number of stochastic factors. The HJM model uses the current yield curve and forward rate curve, and then specifies a continuous time stochastic process to describe the evolution of the yield curve over a specified time period.

The model is summarised here only; readers interested in the derivation of the model are directed to the original paper or a discussion of it in Baxter and Rennie (1996), Hull (1997) or James and Webber (2000). To describe the model we use the following notation:

(t, T)	is the trading interval for a fixed period from t to T, where $t > 0$
$W(t)$	is the independent Brownian motion or Weiner process that describes the interest rate process
(Ω, F, Q)	is the probability space where F is the σ-algebra representing measurable events and Q is the measure of probability
$P(t, T)$	is the price at time t of a zero-coupon bond that matures at time T. The bond has a redemption value of 1 at time T.

2 See Rebonato and Cooper (1996).

3 Heath, D., Jarrow, R., Morton, A., ‘Bond pricing and the term structure of interest rates: a new methodology’, *Econometrica* 60(1), 1992, pp. 77–105.

The instantaneous forward rate $f(t, T)$ at time t is given by (5.26):

$$f(t,\ T) = \frac{-\partial\ ln\ P(t,\ T)}{\partial T} \tag{5.26}$$

and describes the interest rate that is applicable on a default-free loan at time t for the period from T to a point one instant later. In their paper Heath, Jarrow and Morton state that the solution to the differential equation (5.26) results in the expression for the price of the bond, shown in (5.27):

$$P(t,\ T) = exp\left(-\int_t^T f(t,\ s)\mathrm{d}s\right) \tag{5.27}$$

while the spot interest rate at time t is the instantaneous forward rate at time t for maturity date t, shown by:

$$r(t) = f(t,\ t).$$

We now describe the model's exposition of movements in the term structure.

The HJM model describes the evolution of forward rates given an initial forward rate curve, which is taken as given. For the period $T \in [0,\ t]$ the forward rate $f(t, T)$ satisfies the equation (5.28):

$$f(t,\ T) - f(0,\ T) = \int_0^t \alpha(\nu,\ T,\ \omega)\mathrm{d}\nu + \sum_{i=1}^{n}\int_0^t \sigma_i(\nu,\ T,\ \omega)\mathrm{d}W(t). \tag{5.28}$$

The expression describes a stochastic process composed of n independent Weiner processes, from which the whole forward rate curve, from the initial curve at time 0, is derived. Each individual forward rate maturity is a function of a specific volatility coefficient. The volatility values $(\sigma_i(t,\ T,\ \omega))$ are not specified in the model and are dependent on historical Weiner processes. From (5.28) following the HJM model the spot rate stochastic process is given by (5.29):

$$r(t) = f(0,\ t) + \int_0^t \alpha(\nu,\ t,\ \omega)\mathrm{d}\nu + \sum_{i=1}^{n}\int_0^t \sigma_i(\nu,\ t,\ \omega)\mathrm{d}W(t) \tag{5.29}$$

for the period $t \in (0,\ T)$.

The model then goes on to show that the process of changes in the bond price is given by:

$$\begin{aligned} ln\ P(t,\ T) = ln\ P(0,\ T) &+ \int_0^t (r(\nu) + b(\nu,\ T))\mathrm{d}\nu \\ &- \frac{1}{2}\sum_{i=1}^{n}\int_0^t a_i(\nu,\ T)^2\mathrm{d}\nu + \sum_{i=1}^{n}\int_0^t a(\nu,\ T)\mathrm{d}W(t) \end{aligned} \tag{5.30}$$

where $a_i(t,\ T,\ \omega) \equiv -\int_t^T \sigma_i(t,\ \nu,\ \omega)\mathrm{d}\nu \quad \text{for} \quad i = 1,\ 2, \ldots,\ n$ and

$$b(t,\ T,\ \omega) \equiv -\int_t^T \alpha(t,\ \nu,\ \omega)\mathrm{d}\nu + \frac{1}{2}\sum_{i=1}^{n} a_i(t,\ T,\ \omega)^2.$$

The expression in (5.30) describes the dynamics of the bond price as a continuous stochastic process with a drift of $(r,\ (t,\ \omega)) + b(t,\ T,\ \omega)$ and a volatility value of $a_i(t,\ T,\ \omega)$.

The no-arbitrage condition is set by defining the price of a zero-coupon bond that matures at time T in terms of an 'accumulation factor' $B(t)$ which is the value of a money market account that is invested at time 0 and reinvested at time t at an interest rate of $r(t)$. This accumulation factor is defined as (5.31):

$$B(t) = exp\left(\int_0^t r(y)\mathrm{d}y\right) \tag{5.31}$$

and the value of the zero-coupon bond in terms of this accumulation factor is $Z(t, T) = P(t, T)/B(t)$ for the period $T \in (0, t)$ and $t \in (0, T)$.

Following HJM, by applying Itô's lemma the model obtains the following result for $Z(t, T)$, shown in (5.32):

$$ln\, Z(t, T) = ln\, Z(0, T) + \int_0^t b(\nu, T)\mathrm{d}\nu - \frac{1}{2}\sum_{i=1}^n \int_0^t a_i(\nu, T)^2 \mathrm{d}\nu + \sum_{i=1}^n \int_0^t a_i(\nu, T)\mathrm{d}W(t). \tag{5.32}$$

In the HJM model the processes for the bond price and the spot rate are not independent of each other. As an arbitrage-free pricing model it differs in crucial respects from the equilibrium models presented in the previous chapter. The core of the HJM model is that given a current forward rate curve, and a function capturing the dynamics of the forward rate process, it models the entire term structure.

A drawback of the model is that it requires the input of instantaneous forward rates, which cannot necessarily be observed directly in the market. Models have been developed that are in the HJM approach that take this factor into account, including those presented by Brace, Gatarek and Musiela (1997) and Jamshidian (1997). This family of models is known as the LIBOR market model or the BGM model. In the BGM model there is initially one factor, the forward rate $f(t)$ which is the rate applicable from time t_k to time t_{k+1} at time t. The forward rate is described by (5.33):

$$\mathrm{d}f(t) = \theta(t)f(t)\mathrm{d}W \tag{5.33}$$

where the market is assumed to be forward risk-neutral.

The relationship between forward rates and the price of a zero-coupon bond at time t is given by (5.34):

$$\frac{P(t, t_i)}{P(t, t_{i+1})} = 1 + \delta_i f_i(t) \tag{5.34}$$

where δ_i is the compounding period between t and t_{i+1}.

To determine the volatility of the zero-coupon bond price $\nu(t)$ at time t, it can be shown that applying Itô's lemma to (5.34) we obtain

$$\nu_i(t) - \nu_{i+1}(t) = \frac{\delta_i f_i(t)\theta_i(t)}{1 + \delta_i f_i(t)}. \tag{5.35}$$

It is possible to extend the BGM model to incorporate more independent factors.

5.3.2 *Jump-augmented models*

Further research has produced a category of models that attempt to describe the jump feature of asset prices and interest rates. Observation of the markets confirms that many asset price patterns and interest rate changes do not move continuously from one price (rate) to another, but sometimes follow a series of jumps. A good example of a jump movement is when a central bank changes the base interest rate; when this happens, the entire yield curve shifts to incorporate the effect of the new base rate. There is a considerable body of literature on the subject, and we only refer to a small number of texts here.

One type of jump model is the jump-augmented HJM model, described in Jarrow and Madan (1991), Bjork (1996) and Das (1997). This is not described here because we have not covered the necessary technical background. Another is the jump-augmented Vasicek model described by Das and Foresi (1996) and Baz and Das (1996). In this, the short rate process is captured by

$$\mathrm{d}r_t = \alpha(\mu - r_t)\mathrm{d}t + \sigma \mathrm{d}W_t + J\mathrm{d}N_t \tag{5.36}$$

where N_t is a Poisson process with a constant intensity λ and J is a random jump size.

Other jump models have been described by Attari (1996), Das and Foresi (1996) and Honore (1998).

5.4 Assessing one-factor and multi-factor models

In assessing the value of the different models that have been developed and the efficacy of each, what is important is how they can be applied in the market, rather than any notion that multi-factor models are necessarily ‘better’ than one-factor models because they are somehow more ‘real-world’. What is required is a mechanism that efficiently prices bonds and interest-rate options; a term structure model attempts to accomplish this by describing the dynamics of the interest rate process and generating random interest-rate paths. The generated paths are then used to discount the cash flows from the fixed income instrument, having initially been used to generate the cash flows in the first place. In practice a one-factor model that has been accurately calibrated will value fixed income instruments efficiently. This is because of the determinants of bond pricing; to illustrate, consider a fixed income instrument with a fixed maturity date. To value such a bond at a particular time, we need only know the bond yield at that time, and this is essentially a one-factor process. Similarly for a callable bond: when generating its cash flow, we will know whether it will be called by knowing its price at a future date. Generating this cash flow from the interest-rate model is again a one-factor process. Therefore if we are pricing a bond, the dynamics of the price process can usually be adequately described by (5.37):

$$\mathrm{d}P = \mu_P P\mathrm{d}t + \sigma_P P\mathrm{d}W(t) \tag{5.37}$$

which is the process followed by for example the Black–Scholes option model when used to price an interest-rate option. This model does not discount the forward price of the option, which is the second part of the B–S approach: that of assuming a single continuously compounded short-term risk-free interest-rate.

While this approach would work in practice, this would only be for a single security portfolio; it would be unwieldy and inaccurate for valuing a number of securities. As banks and market makers must value many hundreds of cash and off-balance sheet instruments, another approach is required. This other approach was considered in this chapter and involves describing the dynamics of the bond price process in the form of a term structure model. Under this situation a multi-factor model may be more suitable, particularly when used to value options.

To consider one-factor models then, we know that the yield of a bond at a future date is essentially a one-factor process, so a one-factor model may well be accurate. A one-factor model describes only parallel shift yield changes, and it assumes that bond yields and discount rates are perfectly correlated, so that it will not generate all the possible paths of the future discount rate. In practice however much yield curve movement is close to a parallel shift, so this may not be as much of a problem for the majority of situations. If a term structure model accurately reflects the random evolution of the price of a bond, and the actual current rate and forward rate volatilities of the bond are as generated by the model, then the model can be considered effective, and it will generate reasonable cash flow scenarios and with accurate probabilities. It is possible to achieve this with one-factor models. Essentially then, a bank can use a one-factor model when conditions are appropriate, and need only use a multi-factor model where the one-factor model cannot be expected to be accurate. That said, why not simply use a multi-factor model at all times? The main reason is because generating forward rates and valuations from a multi-factor model is a time-consuming process, employing considerable computing power, and as rapidity of analysis and response is of the essence in the markets, it is logical to use a slower model only when it is significantly more accurate than the one-factor model.

It is generally accepted that one-factor models can be used for most bond applications; where multi-factor models are more appropriate may be in the following situations:

- where the instrument being valued is linked to two different interest rates, for example an interest-rate quanto option, or an option with a payoff profile that is a function of the spread between two different reference rates;
- for the valuation of long-dated options or deeply in-the-money or out-of-the-money options, which are affected by the volatility smile. As a stochastic volatility factor will impact the price, a model that assumes constant volatilities would be inaccurate;
- for the valuation of securities that to some extent reflect the slope of the yield curve, such as certain mortgage-backed bonds whose level of prepayment is sometimes a function of the slope of the yield curve;
- for the valuation of very long-dated options, where all possible paths of the future discount rate may be required.

The optimum approach would appear to be a combination of a one-factor model and a multi-factor model to suit individual requirements. However this may not be practical; it might not be ideal to have different parts of a bank using different models (although this does happen; desks across the larger investment banks sometimes use different models) and valuing instruments using different models. The key factors to focus on are accessibility, accuracy, appropriateness and speed of computation.

5.4.1 *Choosing the model*

There are essentially two approaches to modelling the term structure that we have discussed in this and the previous chapter. The Ho–Lee and HJM models begin with the evolution of the whole yield curve, while the BDT, Hull–White and other models specify the dynamics of the short rate, and determine the parameters so that the model itself corresponds to the current term structure. We have also discussed the relative merits of the equilibrium model approach and the no-arbitrage approach. In this final section we discuss the different issues that apply in each case.

Essentially there are two dimensions to consider: risk-neutral versus realistic and equilibrium versus no-arbitrage. There are situations under which each approach may be applied with validity.

No-arbitrage, risk-neutral approach

A commonly encountered approach is the risk-neutral, no-arbitrage model. This is a no-arbitrage model used frequently to value interest-rate options, using parameters that have been interpolated from a set of current market prices rather than estimated from actual historical data. This approach is valid when there is a reliable set of observable market prices and rates. Note that two different no-arbitrage models that are applied in a risk-neutral framework will only generate identical term structures and valuations if exactly identical input parameters have been used. The actual type of model used will have a significant effect on the valuation that is achieved. If however market data is not readily observable, or not reliable, this approach can lead to inaccuracies. This can be expected in illiquid markets such as that for certain long-dated exotic options; where this occurs there is no way to estimate a correlation term structure that allows the model to interpolate between the option prices, because there are two of them or their reliability is not accepted. In this scenario, a multi-factor model that captures the correlations between interest rates of different maturities, as well as the impact of the shape of the yield curve on these correlations, may be more valid. A model with a good statistical fit to the historical correlations recorded by the option product may therefore produce more robust prices. We hesitate to say ‘accurate’ because, in an illiquid market with only a small number of market makers, who is to say which price is the most accurate? For certain exotic options the valuation depends on each market making bank’s valuation model, and how effectively the model has been calibrated.

No-arbitrage, realistic approach

A no-arbitrage model that is implemented in a realistic approach matches precisely the term structure of interest rates that are implied by the current (or initial) observed market yields. It then derives a forward curve for the future that is dependent on the way it has modelled the dynamics of the interest-rate process, which is a measure of probability. This approach is valid when it is important that the initial yield curve must be identical to the current observed yield curve, and is often used for analysing hedging strategy or portfolio strategies. In implementing this method it is sometimes difficult to evaluate the efficacy of the model because of problems in discriminating between model error and exogenous effects. In this approach the model parameters are set to match precisely observed market yields, with no regard to historical data, and there is little degree of freedom by which one can evaluate the model results. Only in a situation where the model generated an identical true term structure, so that the time-dependent parameters resulted in no pricing error at all points along the term structure, and for all dates past and forward, could the model be described for certainty as accurate. Otherwise the difficulty in assessing the effectiveness of this approach means that it is rarely used in practice.

Equilibrium, risk-neutral approach

The second approach is an equilibrium model under risk-neutral conditions. This is also valid under certain conditions. Remember that a no-arbitrage model uses input parameters that are based on observed prices and yields. However observed bond yields reflect a number of factors that frequently distort them away from 'fair value', resulting in a discount function that is also distorted. For instance we saw in our discussions in Part I that bond prices reflect liquidity, benchmark effects, supply and demand and other factors, which can include taxation, coupon size, convexity effect and so on. The same applies in the government zero-coupon bond market. Therefore a no-arbitrage model will use parameters that have been distorted by these factors. However equilibrium models are able to capture the global behaviour of the term structure over a long-term period, which has the effect of stripping out the market distortions (they are treated as 'noise'). Therefore risk-neutral equilibrium models have an advantage over no-arbitrage models as they are not as sensitive to external market factors. In addition when used to price bonds today, equilibrium models can be estimated from historical data when market-observed current prices are unavailable or unreliable. So we conclude that one of the most appropriate times to adopt the risk-neutral equilibrium approach is when observed market yields are not available or subject to excessive distortions. This brings us to the subject of *horizon pricing*, which is the estimation of prices for a bond or other instrument under some expected future market state. While parameters are usually available, and reasonably reliable (in developed markets) for use in current pricing, they will not be available for a scenario-type valuation. In this situation, no-arbitrage models cannot be used at all, because they require the input of a set of market yields, which would not be available for horizon pricing as they would be unknown. Therefore in this case we use an equilibrium model, otherwise no analysis would be possible.

Equilibrium, realistic approach

In fact the inappropriateness of the no-arbitrage realistic approach means that only the equilibrium realistic approach is available. This methodology is used where speed of computation and accuracy of term structure generation and valuation are not of prime importance. It is used most frequently for risk management, regulatory and testing purposes; this includes value-at-risk calculations, VaR stress testing, capital adequacy calculations and other scenario purposes. In this approach and with equilibrium models the derived current term structure that is generated will not match the actual current term structure precisely. This has led to some analysts suggesting that any testing performed with the model will not be perfectly realistic. However the main purpose behind scenario analysis is to assess the impact of different situations; there is nothing illogical about comparing the effect of a theoretical future yield curve on an asset book held today and valued using today's actual yield curve. An equilibrium model is a statistical model of the behaviour of the term structure of rates; therefore using it implies an acceptance that its derived curve will differ from the observed curve. Using a no-arbitrage model would imply that the current term structure model was completely accurate. Therefore for risk management and capital purposes it is common to encounter the equilibrium model, realistic approach.

5.4.2 *Choosing the model: second-time around*

It is important to remain focused on the practical requirements of interest-rate modelling. Market participants are more concerned with the ease with which a model can be implemented, and its accuracy with regard to pricing. In practice different models are suited to different applications, so the range of products traded by a market practitioner will also influence which model is chosen. For instance the extended Vasicek model can be fitted very accurately to the initial term structure, and its implementation is relatively straightforward, being based on a lattice structure. It is also able to accurately price most products, however like all one-factor models it is not a valid model to use when pricing instruments that are sensitive to two or more risk factors, for example quanto options. The extended CIR model is also tractable, although it has a more restricted set of term structures compared to the extended Vasicek model, as a result of the limitations imposed by the $\sqrt{r_t}$ term on the volatility parameter. Both types of models are unable to capture the dynamics of the whole yield curve, for which HJM models must be used.

A drawback of these models is that although they fit the initial term structure, due to their structure they may not continue to calculate prices as the term structure evolves. In practice the models must be re-calibrated frequently to ensure that they continue to describe term structure volatilities that exist in the market.

In selecting the model, a practitioner will select the market variables that are incorporated in the model; these can be directly observed such as zero-coupon rates or forward rates, or swap rates, or they can be indeterminate such as the mean of the short rate. The practitioner will then decide the dynamics of these market or *state variables*, so for example the short rate may be assumed to be mean reverting. Finally the model must be calibrated to market prices, so the model parameter values input must be those that produce market prices as accurately as possible. There are a number of ways that parameters can be estimated; the most common techniques of calibrating to time series data such as interest rate data are *general method of moments* and the *maximum likelihood* method. For information on these estimation methods refer to the bibliography.

Models exhibit different levels of sensitivity to changes in market prices and rates. The extent of a model's sensitivity will also influence the frequency with which the model must be re-calibrated. For example the Black–Derman–Toy model is very sensitive to changes in market prices; because it is a log-r model changes in the process of the underlying variable are larger as they are log-r, than those in the process for rt itself. Some practitioners believe that as they take bond prices and the term structure as given, arbitrage models suffer from an inherent weakness. Liquidity and other considerations frequently result in discrepancies between market yields and theoretical value, and such discrepancies would feed through into an arbitrage model. This drawback of arbitrage models means that users must take care about term structure inputs, and the curve fitting techniques and *smoothing* techniques that are used become critical to model effectiveness. This is discussed in the next chapter.

Other considerations are detailed below.

- **Model inputs:** Arbitrage models use the term structure of spot rate as an input, and this data is straightforward to obtain. Equilibrium models require a measure of the investor's market risk premium, which is rather more problematic. Practitioners analyse historical data on interest rate movements, which is considered less desirable.
- **Using models as part of bond trading strategy:** A key element of market makers' and proprietary traders' strategy is relative value trading, which includes simultaneous buying and selling of certain bonds against others, or classes of bonds against other classes. A yield curve spread trade is a typical relative value trade. How does one

determine relative value?[4] Using an interest-rate model is the answer. For such purposes though, only equilibrium models can be used. By definition since arbitrage models take bond prices and the current term structure as given, they clearly cannot be used to assess relative value. This is because the current price structure would be assumed to be correct. If one were to use such a model for a yield curve trade, it would imply a zero profit potential. Therefore only equilibrium models can be used for such purposes.

- **Model consistency:** As we have noted elsewhere, using models requires their constant calibration and re-calibration over time. For instance, an arbitrage model makes a number of assumptions about the interest rate drift rate and volatility, and in some cases the mean reversion of the dynamics of the rate process. Of course these values will fluctuate constantly over time, so that the estimate of these model parameters used one day will not remain the same over time. So the model will be inconsistent over time and must be re-calibrated to the market. Equilibrium models use parameters that are estimated from historical data, and so there is no unused daily change. Model parameters remain stable. Over time therefore these models remain consistent, at least with themselves. However given the points we have noted above, market participants usually prefer to use arbitrage models and re-calibrate them frequently.

We have only touched on the range of considerations that must be followed when evaluating and implementing an interest-rate model. This is a complex subject with a number of factors to consider, and ongoing research in the area serves to reinforce the fact that it is an important and very current topic.

Selected bibliography and references

Attari, M., 'Discontinuous interest rate processes: an equilibrium model for bond option prices', Working Paper, University of Iowa, 1996, pp. 1–32.

Baxter, M., Rennie, A., *Financial Calculus*, Cambridge University Press, 1996.

Baxter, M., Rennie, A., *An Introduction to Derivative Pricing*, Cambridge University Press, 1998, Chapter 5.

Baz, J., Das, S., 'Analytical approximations of the term structure for jump diffusion processes: a numerical analysis', *Journal of Fixed Income*, June 1996, pp. 78–86.

Bjork, T., 'Interest rate theory', *CIME Lectures*, Working Paper, Stockholm School of Economics, 1996, pp. 1–90.

Black, F., Karasinski, P., 'Bond and option pricing when short rates are lognormal', *Financial Analysts Journal*, Jul–Aug 1991, pp. 52–59.

Brace, A., Gatarek, D., Musiela, M., 'The market model of interest-rate dynamics', *Mathematical Finance* 7(2), 1997, pp. 127–155.

Brennan, M., Schwartz, E., 'An equilibrium model of bond pricing and a test of market efficiency', *Journal of Financial and Quantitative Analysis* 17(3), September 1982, pp. 301–329.

Cheyette, O., 'Term structure dynamics and mortgage valuation', *Journal of Fixed Income*, March 1992, pp. 28–41.

[4] Some traders determine relative value by conducting the analysis inside their head! Nowadays one needs to back up one's gut feeling with formal analysis. A term structure model will assist.

Das, S., 'A direct discrete-time approach to Poisson–Gaussian bond option pricing in the Heath–Jarrow–Morton model', Working Paper, Harvard Business School, 1997, pp. 1–44.

Das, S., Foresi, S., 'Exact solutions for bond and option prices with systematic jump risk', *Review of Derivatives Research* 1, 1996, pp. 1–24.

Duffie, D., Kan, R., 'A yield-factor model of interest rates', *Mathematical Finance* 6(4), 1996, pp. 379–406.

Fitton, P., McNatt, J., 'The four faces of an interest rate model', in Fabozzi, F. (ed.), *Advances in Fixed Income Valuation Modelling and Risk Management*, FJF Associates, 1997.

Harrison, M., Pliska, S., 'Martingales and stochastic integrals in the theory of continuous trading', *Stochastic Processes and their Applications II*, 1981, pp. 215–260.

Heath, D., Jarrow, R., Morton, A., 'Bond pricing and the term structure of interest rates: a discrete time approximation', *Journal of Financial and Quantitative Analysis* 25(4), December 1990, pp. 419–440.

Heath, D., Jarrow, R., Morton, A., 'Bond pricing and the term structure of interest rates: a new methodology', *Econometrica* 60(1), 1992, pp. 77–105.

Ho, T.S.Y., Lee, S-B., 'Term structure movements and pricing interest rate contengent claims', *Journal of Finance* 41, 1986, pp. 1011–1029.

Honore, P., 'Five essays on financial econometrics in continuous-time models', PhD Thesis, Aarhus School of Business, 1998.

Hughston, L. (ed.), *Vasicek and Beyond: Approaches to Building and Applying Interest Rate Models*, Risk Publications, 1996.

Hull, J., *Options, Futures and Other Derivatives*, Wiley, 1997.

Hull, J., White, A., 'Bond option pricing based on a model for the evolution of bond prices', *Advances in Futures and Options Research* 6, 1993, pp. 1–13.

Hull, J., White, A., 'Numerical procedures for implementing term structure models II: two-factor models', *Journal of Derivatives* 2(2), Winter 1994, pp. 37–48.

James, J., Webber, N., *Interest Rate Modelling*, Wiley, 1997.

Jamshidian, F., 'LIBOR and swap market models and measures', *Finance and Stochastics* 1, 1997, pp. 293–330.

Jarrow, R., Madan, D., 'Option pricing using the term structure of interest rates to hedge systematic discontinuities in asset returns', Working Paper, Cornell University, 1991, pp. 1–47.

Jarrow, R., Turnbull, S., *Derivative Securities*, South-Western Publishing, 1996.

Rebonato, R., *Interest Rate Option Models*, 2nd edition, Wiley, 1998.

Rebonato, R., Cooper, I., 'The limitations of simple two-factor interest-rate models', *Journal of Financial Engineering*, March 1996.

Ritchken, P., Sankarasubramanian, L., 'Volatility structures of forward rates and the dynamics of the term structure', *Mathematical Finance* 5, 1995, pp. 55–72.

Uhrig, M., Walter, U., 'A new numerical approach for fitting the initial yield curve', *Journal of Fixed Income*, March 1996, pp. 82–90.

References on estimation methods

Andersen, T., Lund, J., 'Estimating continuous-time stochastic volatility models of the short-term interest rate', *Journal of Econometrics* 77, 1997, pp. 343–377.

Brown, R., Schaefer, S., 'The term structure of real interest rates and the Cox, Ingersoll and Ross model', *Journal of Financial Economics* 35, 1994, pp. 2–42.

Campbell, J., Lo, A., MacKinlay, C., *The Econometrics of Financial Markets*, Princeton University Press, 1997.

Davidson, R., MacKinnon, J., *Estimation and Inference in Econometrics*, Oxford University Press, 1993.

Duffie, D., Singleton, K., 'Simulated moments estimation of Markov models of asset prices', *Econometrica* 61, 1993, pp. 929–952.

Hansen, L., 'Large sample properties of generalized method of moments estimators', *Econometrica* 50, 1982, pp. 1029–1055.

Lo, A., 'Statistical tests of contingent claims asset pricing models', *Journal of Financial Economics* 17, 1986, pp. 143–173.

Longstaff, F., Schwartz, E., 'Interest rate volatility and the term structure: a two-factor general equilibrium model', *Journal of Finance* 47, 1992, pp. 1259–1282.

6 Estimating and Fitting the Term Structure I

The two previous chapters introduced and described a fraction of the most important research into interest rate models that has been carried out since the first model, presented by Oldrich Vasicek, appeared in 1977. These models can be used to price derivative securities, and equilibrium models can be used to assess fair value in the bond market. Before this can take place however a model must be fitted to the yield curve, or *calibrated*.[1] In practice this is carried out in two ways; the most popular approach involves calibrating the model against market interest rates given by instruments such as cash Libor deposits, futures, swaps and bonds. The alternative method is to model the yield curve from the market rates and then calibrate the model to this fitted yield curve. The first approach is common when using for example extended Vasicek models, while the second technique is more useful with whole yield curve models such as the Heath–Jarrow–Morton model.

There are a number of techniques that can be used to fit the yield curve. These include regression methods and *spline* techniques. More recent methods such as *kernel approximations* and *linear programming* are also beginning to be used by practitioners. In this chapter we provide an introduction to some of these, however a detailed exposition would warrant a book in its own right. We discuss fitting the spot and forward yield curve and review the methods used to estimate spot and forward yield curves. We then illustrate the cubic spline method for fitting a yield curve from observed government bond yields. There is a large body of literature on this subject. For further information readers are recommended to review Anderson *et al.* (1996) and James and Webber (2000) for the most important research, and interested readers may also wish to consider Bliss (1997), Dahlquist and Svensson (1996) and Waggoner (1997). Alternative approaches are given in Kim (1993) and Zheng (1994).

For a number of reasons practitioners, investors, central banks and government authorities are interested in fitting the zero-coupon yield curve, or the true term structure of interest rates. The use of yield curves is standard in monetary policy analysis, and central banks are increasingly making use of forward interest rates for this purpose as well. Forward rates must be estimated from the yield curve that has been constructed from current market yields, generally T-bill and government bond yields. Particularly useful information that can be derived from government bond prices includes the yield curve for implied forward rates, as these reflect the market's expectations of the future path of interest rates.[2] They are also

[1] In fact a model needs to be calibrated to the market, but the most important item against which it must be calibrated is the current term structure.

[2] Remember of course that the forward rate is derived from the current spot rate term structure, and therefore although it is an expectation based on all currently known information, it is not a prediction of the term structure in the future. Nevertheless the forward rate is important because it enables market makers to price and hedge financial instruments.

used by the market to price bonds and determine the extent of the credit spread applicable to corporate bonds. The requirements of the monetary authorities however are slightly different to those of market practitioners: central bankers and the government are not so concerned with the accuracy of the spot curve with regard to pricing securities; rather, they are interested in the information content of the fitted curve, particularly concerning implied forward rates and the market expectations of future interest rates and levels of inflation. In the second part of this chapter we review the information content of the yield curve in the UK gilt market.

6.1 Introduction

From an elementary understanding of the markets we know that there is a relationship between a set of discount factors, and the discount function, the par yield curve, the zero-coupon yield curve and the forward yield curve. If we know one of these functions we may readily compute the other three. In practice although the zero-coupon yield curve is directly observable from the yields of zero-coupon government bonds, liquidity and investor preferences usually mean that a theoretical set of all these curves is derived from the yields of coupon government bonds in the market. There are a number of ways that the zero-coupon curve can be fitted, using either a discount function or the par yield curve.

The pricing of financial instruments in the debt market revolves around the yield curve. The use in the market and by the central authorities of the government term structure to ascertain the market's expectation of future interest rates is well established. This reflects the fact that the spot yield curve is the geometric average of the same maturity structure implied forward rates. Here we discuss the information content of the yield curve, and how the zero-coupon curve may be best fitted to enable analysts to extract information from the implied forward rate yield curve. This is used for a number of purposes by central government monetary authorities and by analysts and economists. In the United Kingdom for example the yield on government bonds is used as the benchmark for interest charges to local authorities and public sector bodies. Yield curve data may also be used as one of the parameters for a general interest rate model (for example, see Cox, Ingersoll, Ross (1985) for a one-factor model and Heath, Jarrow, Morton (1992) for a multi-factor model).

Although the use of yield curves is quite common as part of monetary policy analysis, central banks such as the US Federal Reserve and the Bank of England have only recently begun to use *forward* interest rates as an indicator for monetary policy purposes. We know that a forward rate is an interest rate applicable to a debt instrument whose term begins at a future date, and ends at a date beyond that. Although there is a market in forward rates, the prices at which forward instruments are quoted are derived from spot interest rates. That is, *implied* forward rates are calculated from the spot yield curve, which is in turn modelled from the prices of instruments in the market, usually government bills and bonds. This implies that the shape and position of the spot curve reflects market belief on future interest rates, which is why it is used to calculate forward rates. The information content and predictive power of a spot term structure is based on this belief. Forward rates may be estimated using any one of a number of models. They can be interpreted as reflecting the market's expectations of future *short-term* interest rates, which in turn are indicators of expected inflation levels. The same information is contained in the spot yield curve, however monetary authorities often prefer to use forward rates as they are better applied to policy analysis. Whereas the spot yield curve is the expected average of forward rates, the forward rate curve reflects the expected future time period of one short-term forward rate. This means that the forward curve can be split into short-term and long-term segments in a more straightforward fashion than the spot curve.

As it is used as a predictive indicator, the spot yield curve needs to be fitted as accurately as possible. This is an area that has been extensively researched (see McCulloch 1975, Deacon and Derry 1994, Schaefer 1981, Waggoner 1997, Nelson and Siegel 1987, Svensson 1994, 1995 *inter alia*). Invariably researchers use the government debt market as the basis for modelling the term structure. This is because the government market is the most liquid debt market in any country, and also because (in a developed economy) government securities are default-free, so that government borrowing rates are considered risk-free. Whatever method is used to fit the term structure, it should aim to meet the following criteria when the main use is for government policy, rather than the pricing of financial instruments:

- the method should attempt to fit implied forward rates, because the primary objective is to derive the forward curve and not market spot yields;
- the resulting derived forward curve should be as smooth as possible, again because the aim is to provide information on the future level and direction of interest rates, and expectations on central bank monetary policy, rather than an accurate valuation of financial instruments along the maturity term structure;
- it should have as few market assumptions as possible.

There are a number of curve fitting methods that may be employed. In the United Kingdom gilt market the Bank of England previously used an in-house model[3] but has since adopted a modified technique proposed by Svensson (1995) and subsequently Fisher, Nychka and Zervos (1995), Waggoner (1997) and Anderson and Sleath (1999). The Waggoner method is discussed in a later section. In the UK the introduction of a market in government zero-coupon bonds has enabled the accuracy of a fitted spot term structure to be compared to actual market spot rates; there is also useful information to be gleaned from using data from the gilt repo market when comparing the accuracy of the short-end of the fitted curve, as discussed by Anderson and Sleath (1999). We set the scene below.

6.2 Bond market information

6.2.1 *Basic concepts*

Central banks and market practitioners use interest rates prevailing in the government bond market to extract certain information, the most important of which is implied forward rates. These are an estimate of the market's expectations about the future direction of short-term interest rates. They are important because they signify the market's expectations about the future path of interest rates, however they are also used in derivative pricing and to create synthetic bond prices from the extent of credit spreads of corporate bonds.

3 See Mastronikola (1991).

Forward rates may be calculated using the discount function or spot interest rates. If spot interest rates are known then the bond price equation can be set as:

$$P = \frac{C}{(1+rs_1)} + \frac{C}{(1+rs_2)^2} + \cdots + \frac{C+M}{(1+rs_n)^n} \tag{6.1}$$

where

C is the coupon
M is the redemption payment on maturity (par)
rs_t is the *spot interest rate* applicable to the cash flow in period t $(t = 1, \ldots, n)$.

The bond price equation is usually given in terms of *discount factors*, with the present value of each coupon payment and the maturity payment being the product of multiplying them by their relevant discount factors. This allows us to set the price equation as shown by (6.2),

$$P = \sum_{t=1}^{n} C\mathrm{d}f_t + M\mathrm{d}f_n \tag{6.2}$$

where $\mathrm{d}f_t$ is the t-period discount factor $(t = 1, \ldots, m)$ given by (6.3):

$$\mathrm{d}f_t = \frac{1}{(1+rs_t)^t}, \quad t = 1, \ldots, m. \tag{6.3}$$

A discount factor is a value for a discrete point in time, whereas markets often prefer to think of a continuous value of discount factors that applies a specific discount factor to any time t. This is known as the *discount function*, which is the continuous set of discrete discount factors and is indicated by $\mathrm{d}f_t = \delta(t_t)$.

The discount function relates the current cash bond yield curve with the spot yield curve and the implied forward rate yield curve. From (6.3) we can set:

$$\mathrm{d}f_t = (1+rs_t)^{-t}.$$

As the spot rate rs is the average of the implied short-term forward rates $rf_1, rf_2, \ldots, rf_t$ we state

$$\begin{aligned} 1/\mathrm{d}f_1 &= (1+rs_1) = (1+rf_1) \\ 1/\mathrm{d}f_2 &= (1+rs_2)^2 = (1+rf_1)(1+rf_2) \\ 1/\mathrm{d}f_t &= (1+rs_t)^t = (1+rf_1)(1+rf_2)\cdots(1+rf_t). \end{aligned} \tag{6.4}$$

From (6.4) we see that $1 + rs_t$ is the geometric mean of $(1+rf_1), (1+rf_2), \ldots, (1+rf_t)$.

Implied forward rates indicate the expected short-term (one-period) future interest rate for a specific point along the term structure; they reflect the spread on the marginal rate of return that the market requires if it is investing in debt instruments of longer and longer maturities.

In order to calculate the range of implied forward rates we require the term structure of spot rates for all periods along the continuous discount function. This is not possible in practice because a bond market will only contain a finite number of coupon-bearing bonds maturing on discrete dates. While the coupon yield curve can be observed, we are then required to ‘fit’ the observed curve to a continuous term structure. Note that in the UK gilt market for example, there is a zero-coupon bond market, so that it is possible to observe spot rates directly, but for reasons of liquidity, analysts prefer to use a fitted yield curve (the *theoretical* curve) and compare this to the observed curve.

6.2.2 *Estimating yield curve functions*

The traditional approach to yield curve fitting involves the calculation of a set of discount factors from market interest rates. From this a spot yield curve can be estimated. The market data can be money market interest rates, futures and swap rates and bond yields. In general though this approach tends to produce 'ragged' spot rates and a forward rate curve with pronounced jagged knot points, due to the scarcity of data along the maturity structure.[4] A refinement of this technique is to use polynomial approximation to the yield curve.

The McCulloch method (1971, 1975) describes the discount function as a linear combination of a specified number of approximating functions, so for example if there are k such functions on which j coefficients are estimated, the discount function that is generated by the set of approximations is a kth degree polynomial. The drawback of this approach is that unless the market observations are spaced at equal intervals through the maturity range, such a polynomial will fit the long end of the curve fairly inaccurately. To account for this McCulloch proposed using piecewise polynomial functions or splines to approximate the discount function. A polynomial spline can be thought of as a number of separate polynomial functions, joined smoothly at a number of join, break or *knot* points. In mathematics the term 'smooth' has a precise meaning, but in the context of a piecewise r-degree spline it is generally taken to mean that the $(r-1)$th derivative of the functions either side of each knot point are continuous. McCulloch originally used a quadratic spline to estimate the discount function. This results however in extreme bumps or 'knuckles' in the corresponding forward rate curve, which makes the curve unsuitable for policy analysis. To avoid this effect, it is necessary to increase the number of estimating functions and to use a *cubic spline*. This was presented by McCulloch in his second paper, and his specification is summarised in Appendix 6.1.

One of the main criticisms of cubic and polynomial functions is that they produce forward rate curves that exhibit unrealistic properties at the long end, usually a steep fall or rise in the curve. A method proposed by Vasicek and Fong (1982) avoids this feature, and produces smoother forward curves. Their approach characterises the discount function as exponential in shape, which is why splines, being polynomials, do not provide a good fit to the discount function, as they have a different curvature to exponential functions. Vasicek and Fong instead propose a transform to the argument T of the discount function $\nu(T)$. This transform is given by

$$T = -(1/\alpha)ln(1-x), \text{ where } 0 \leq x \leq 1 \tag{6.5}$$

and has the effect of transforming the discount function from an approximately exponential function of T to an approximately linear function of x. Polynomial splines can then be employed to estimate this transformed discount function. Using this transform it is straightforward to impose additional constraints on the discount function. The parameter α constitutes the limiting value of the forward rates, and can be fitted to the date as part of

[4] For a good account of why this approach is not satisfactory see James and Webber (2000, Chapter 15).

the estimation. Vasicek and Fong use a cubic spline to estimate the transformed discount function. In terms of the original variable T this is equivalent to estimating the discount function by a third-order exponential spline, that is between each pair of knot points $\nu(T)$ takes the form:

$$\nu(T) = b_0 + b_1 e^{-2\alpha T} + b_3 e^{3\alpha T}. \tag{6.6}$$

However, Shea (1985) has indicated that in practice exponential splines do not produce more stable estimates of the term structure than polynomial splines. He also recommended using basis splines or *B-splines*, functions that are identically zero over most of the approximation space, to prevent loss of accuracy due to the lack of observations at the long end of the curve.

6.3 Curve-fitting techniques: parametric

There are a number of models that one may use to fit the spot rate term structure. One-factor models (see Vasicek 1977, Dothan 1978, Cox, Ingersoll, Ross 1985) model the one-period short rate to obtain a forward yield curve. The simplest method uses a binomial model of probabilities to model the forward rate. Multi-factor models (see Heath, Jarrow, Morton 1992) express analytically the entire yield curve in terms of two forward rates (or 'spanning' rates). For an analysis of the information content of both methodologies see Edmister and Madan (1993), who conclude that the multi-factor models provide more accurate results. Essentially the information content of the yield curve is best estimated using a multi-factor model, and is more accurate at the longer end of the curve whatever methodology is used. Edmister and Madan also conclude that modelling the short-end of the curve suffers from distortions resulting from government intervention in short-term interest rates.

The Bank of England uses a variation of the Svensson yield curve model, a one-dimensional *parametric* yield curve model. This is similar to the Nelson and Siegel model and defines the forward rate curve $f(m)$ as a function of a set of unknown parameters, which are related to the short-term interest rate and the slope of the yield curve. The model is summarised in Appendix 6.2. Anderson and Sleath (1999) assess parametric models, including the Svensson model, against spline-based methods such as those described by Waggoner (1997), and we summarise their results later in this chapter.

6.3.1 *Parametric techniques*

Curve-fitting techniques generally fit into two classes, as described for example in Chapter 15 of James and Webber (2000), *parametric* methods and *spline-based* methods. Parametric techniques are so-called because they model the forward rate using a parametric function. An early parametric technique was that described by Nelson and Siegel (1987), which models the forward rate curve. Given the relationship between spot and forward rates, such an approach is identical to modelling a spot rate curve by taking a geometric average of the forward rates curve. A fairly flexible function for the forward rate is described in the Nelson-Siegel approach, known as a Laguerre function (plus a constant) and is given by

$$f(T) = \beta_0 + \beta_1 e^{-T/\tau_1} + \frac{\beta_2}{\tau_1} T e^{-T/\tau_1} \tag{6.7}$$

where T is the variable being calculated and β_0, β_1, β_2, and τ_1 are the parameters required to be estimated. Remembering that the spot rate is an average of the forward rates, that is

$$rs = \frac{\int_0^T f(u)\mathrm{d}u}{T} \tag{6.8}$$

then (6.7) implies that the spot rate is given by

$$rs(T) = \beta_0 + (\beta_1 + \beta_2)\frac{\tau_1}{T}(1 - e^{-T/\tau_1}) - \beta_2 e^{-T/\tau_1}. \tag{6.9}$$

To illustrate implementation, we adapt with permission the Anderson and Sleath (1999) evaluation of the Nelson–Siegel method; we set parameter values of:

$$\beta_0 = 5.0 \qquad \beta_1 = -1 \qquad \tau_1 = 1$$

and denote the remaining parameter as a, which reduces (6.9) to

$$rs(T) = 5 + (-1 + a)\frac{1 - e^{-T}}{T} - ae^{-T}. \tag{6.10}$$

Setting β_0 as 5.0 means that the spot rate has been set to a common value of 5.0%. As an exercise we evaluate the possible results with the same parameters used by Anderson and Sleath in their analysis with the exception of the initial spot rate, and change the values for the term to maturity to 10, 20, 30 and 1000 years and the value of a to -5, -3, -1, 0, 1, 3 and 5. Our results are given in Table 6.1. As the value for T increases to very high values the convergence of spot rates to the initial value proceeds only slowly. However, our results illustrate the process.

An evaluation fitting the Nelson–Siegel curve to actual gilt yields from June 1997 is described in the next section.

Another parametric method is described by Svensson (1994, 1995). This adds an extra coefficient to the Nelson–Siegel model and has been described as an extended Nelson–Siegel model. The extra parameter introduces greater flexibility, so that the resulting curve can model forward curves that have more than one 'hump'. It is given by (6.11):

$$f(T) = \beta_0 + \beta_1 e^{-T/\tau_1} + \beta_2 \frac{T}{\tau_1} e^{-T/\tau_1} + \beta_3 \frac{T}{\tau_1} e^{-T/\tau_2}. \tag{6.11}$$

In the Svensson model there are six coefficients β_0, β_1, β_2, β_3, τ_1, τ_2 that must be estimated. The model was adopted by central monetary authorities such as the Swedish Riksbank and the Bank of England (who subsequently adopted a modified version of this model, which we describe shortly, following the publication of the Waggoner paper by the Federal Reserve Bank of England). In their 1999 paper Anderson and Sleath evaluate the two parametric

	***a* values**						
Maturity (*T*) years	**−5**	**−3**	**−1**	**0**	**1**	**3**	**5**
10	4.4003	4.6002	4.8001	4.9	5.0000	5.1999	5.3998
20	4.7000	4.8000	4.9000	4.9500	5.0000	5.1000	5.2000
30	4.8000	4.8667	4.9333	4.9667	5.0000	5.0667	5.1333
1000	4.9940	4.996	4.9980	4.9990	5.0000	5.0200	5.0040

Table 6.1: Spot rate values using Nelson–Siegel model and user-specified parameters.

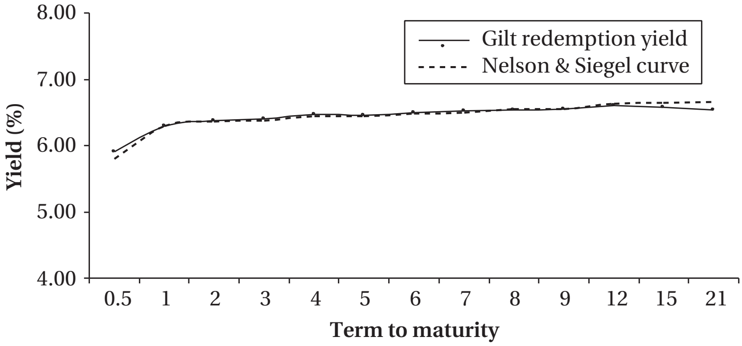


Figure 6.1: A Nelson and Siegel fitted yield curve and gilt redemption yield curve.

techniques we have described, in an effort to improve their flexibility, based on the spline methods presented by Fisher, Nychka and Zervos (1995) and Waggoner (1997).

6.3.2 *Parameterised yield curves*

The technique for curve fitting presented by Nelson and Siegel and variants on it described by Svensson (1994), Wiseman (1994) and Bjork and Christensen (1997) have a small number of parameters, and generally one obtains a relatively close approximation to the yield curve with them. As we saw above, the Nelson and Siegel curve contains four parameters while the Svensson curve has six parameters. The curve presented by Wiseman contains $2 \times (n+1)$ parameters, given by $\{\beta_j, k_j\}_{j=0,\ldots,n}$. The curve is $f_2(\tau)$:

$$f_2(\tau) = \sum_{j=0}^{n} \beta_j e^{-k_j \tau}. \tag{6.12}$$

The original Nelson and Siegel curve does not produce close approximations for all types of yield curves, because the small number of parameters limits flexibility. It can be used to model the spot rate or the forward rate curve, but does not produce accurate results if used to model the discount curve. An example of a fitted Nelson and Siegel curve is shown in Figure 6.1 for UK gilt yields from June 1997. The table of actual gilt yields is shown as well (Table 6.2).

The fitted curve is a close approximation to the redemption yield curve, and is also very smooth. However the fit is inaccurate at the very short end, indicating an underpriced six-month bond, and also does not approximate the long end of the curve. For this reason B-spline methods are more commonly used.

6.4 The cubic spline method for estimating and fitting the yield curve

In mathematical applications a *spline* is *piecewise polynomial*, this being a function that is composed of a number of individual polynomial segments that are joined at user-specified points known as *knot points*. The function is twice-differentiable at each knot point, which produces a smooth curve at each connecting knot point. The commonest approach uses regression methods to fit the spline function, and an excellent and accessible account of this technique is given in Suits *et al.* (1978); the article is summarised in Choudhry (2001).

Term to maturity	Gilt redemption yield %	Term to maturity	Gilt redemption yield %
0.5	5.90	7	6.52
1	6.29	8	6.54
2	6.37	9	6.55
3	6.40	12	6.60
4	6.47	15	6.58
5	6.45	21	6.54
6	6.50		

Table 6.2: Gilt redemption yields. Source: Butler Gilts.

In this section we summarise with permission the spline approach described by Waggoner in his ground-breaking article published by the Federal Reserve Bank of Atlanta in 1997.

Spline methods are commonly used to derive spot and forward rate curves and the discount function from the observed yields of bonds in the market. A popular method is that proposed by McCulloch (1975), which uses regression cubic splines to derive the discount function. Waggoner (1996) has written that this method however, while accurate and stable, produces forward rate curves that oscillate. In fact this property is exhibited by virtually all curve fitting techniques, but the objective for the analyst is to produce curves with the smallest amount of oscillation. A technique posited by Fisher, Nychka and Zervos (1995) used a cubic spline that incorporated a 'roughness penalty' when extracting the forward rate curve. This approach produces a decreased level of oscillation but also reduces the fit of the curve to the actual observed yields. A later technique modified this method by using a 'variable roughness penalty' (Waggoner 1997) and this approach is described here.

6.4.1 *Using a cubic spline: the Waggoner model*

A cubic spline approach can be used as the functional form for the discount function or the forward rate curve. We define a function g on the interval $[t_1, t_N]$ as a cubic spline with node points $t_1 < t_2 < \ldots < t_n$ if it is a cubic polynomial on each of the subintervals $[t_{j-1}, t_j]$ for $1 < j < n$ and if it can be continuously differentiated over the interval $[t_1, t_N]$. The node points are $\tau_1 < \tau_2 < \ldots < \tau_N$ which are the cash flow and maturity dates of the set of bonds (assuming the bonds are semi-annual coupon instruments). Following Waggoner (1997) we set $\tau_0 \equiv 0$ so that the curve is derived from the point zero to the point of the longest-dated bond in the sample. It is possible to use all the node points in the interval to produce the yield curve, however the more points there are in a cubic spline, the greater the tendency for the derived forward curve to oscillate, more so at longer maturities. We wish to minimise the level of oscillation, because for monetary policy purposes the curve is used to provide information on expected future interest rates. A fluctuating yield curve would imply oscillations in expected future prices, and this can produce illogical results, particularly at the long end of the curve. For example, a curve may imply that while the current yield of a six-month T-bill is £97.50, the price of a six-month bill in one year's time will be £98, while the price of such a bill in two years' time will be £95. This is not an unreasonable expectation. However the same implications for six-month bill prices in 25, 26 and 27 years' time is less reasonable.

Therefore the fitted curve should smooth out the forward rates at longer maturities, which calls for a reduced level of oscillation. The McCulloch technique uses regression splines to reduce forward rate fluctuation, while the Fisher *et al.* and the Waggoner approach use a smoothed spline and a modified smoothed spline.

In a regression spline a smaller number of node points are used in order to reduce the level of oscillation. This affects the flexibility of the cubic spline over the interval that is being considered; there is a trade-off between accuracy and the level of oscillation. By reducing node points at the longer end but keeping more at the short end, oscillation is reduced but the curve retains accuracy at the short end. In practice it is common for node points to be set in one of the ways shown in Figure 6.2, but obviously there are any number of ways that node points may be set.

Once we have chosen the node points we set the yield curve ψ as the cubic spline that minimises the function (6.13):

$$\sum_{i=1}^{N}\left(P_i - \hat{P}_i(\psi)\right)^2. \tag{6.13}$$

The technique proposed by McCulloch (1975) used a regression cubic spline to approximate the discount function, and he suggested that the number of node points that are used be roughly equal to the square root of the number of bonds in the sample, with equal spacing so that an equal number of bonds mature between adjacent nodes. A number of writers have suggested that this approach produces accurate results in practice.[5] The discount function is constrained to set $\nu(0) = 1$. Given these parameters the discount function chosen is the one that minimises the function (6.14). As this is a discount function and not a yield curve, (6.14) can be solved using the least squares method.

$$\sum_{i=1}^{N}\left(P_i - \hat{P}_i^{\nu}(\nu)\right)^2. \tag{6.14}$$

For a smoothed spline, the level of oscillation is controlled by setting a ‘roughness penalty’ in the function, and not by reducing the number of node points. The yield curve ψ is chosen that minimises the objective function (6.15):

$$\sum_{i=1}^{N}\left(P_i - \hat{P}_i^{*}(\psi)\right)^2 + \lambda \int_0^{t_N} (\psi''(t))^2 \mathrm{d}t \tag{6.15}$$

for all the cubic splines over the node points $\tau_0 < \tau_1 < \tau_2 < \ldots < \tau_N$. In minimising this function there is a trade-off between the goodness of fit, which is given by the first term, and the degree of smoothness, which is measured by the second term. This trade-off is known as

1w	1m	3m	6m	9m	1y	2y	3y	4y	5y	6y	7y	8y	9y	10y

3m	6m	1y	2y	3y	4y	5y	7y	10y	15y	20y	25y	30y

Figure 6.2: Suggested node points.

5 For example see Bliss (1997).

the 'roughness penalty' and is given by λ, which is a positive constant. If λ is set to zero the function reverts to a regression spline, and as it increases g approaches a linear function. The flexibility of the spline is a function of both the spacing between the node points and the magnitude of λ, although as λ increases the impact of the node spacing decreases. For large values of λ the flexibility of the spline is essentially similar across all terms. This is not necessarily ideal because as we saw from Figure 6.1 we require the spline to be more flexible at the short end, and less so at the long end. Therefore Waggoner (1997) has proposed a modified smoothed spline. For a modified smoothed spline the objective function (6.16) is minimised over the whole term covering the node points $\tau_0 < \tau_1 < \tau_2 < \ldots < \tau_N$.

$$\sum_{i=1}^{N}\left(P_i - \hat{P}_i^*(\psi)\right)^2 + \int_0^{\tau_N} \lambda(t)(\psi''(t))^2 \mathrm{d}t. \tag{6.16}$$

The approach used by Fisher *et al.* (1995) is a smoothed cubic spline that approximates the forward curve. The number of nodes to use is recommended as approximately one-third of the number of bonds used in the sample, spaced apart so that there is an equal number of bonds maturing between adjacent nodes. This is different to the theoretical approach, which is to have node points at every interval where there is a bond cash flow, however in practice using the smaller number of nodes as proposed by Fisher *et al.* produces essentially an identical forward rate curve, but with fewer calculations required. The resulting forward rate curve is the cubic spline that minimises the function (6.17):

$$\sum_{i=1}^{N}\left(P_i - \hat{P}_i^f(f)\right)^2 + \lambda \int_0^{\tau_N} (f''(s))^2 \mathrm{d}s. \tag{6.17}$$

The value of λ is obtained by a method known as generalised cross-validation (GCV). It is the value that minimises the expression in (6.18):

$$\gamma(\lambda) = \frac{rss(\lambda)}{(N - \theta \mathrm{ep}(\lambda))^2} \tag{6.18}$$

where

N	is the number of bonds in the sample
$rss(\lambda)$	is the residual sum of squares, given by $rss(\lambda) = \sum_{i=1}^{N}(P_i - \hat{P}_i^f(f_\lambda))^2$ where f_λ is the forward rate curve that minimises the expression
$\mathrm{ep}(\lambda)$	is the effective number of parameters
θ	is the cost or tuning parameter.

The higher the value for θ, the more rigid is the resulting spline. Fisher *et al.* and Waggoner both set θ equal to 2. Expression (6.17), for a fixed term λ can be solved using a non-linear least squares method. The GCV method can be implemented by using a method known as a *line search*.

Following Fisher *et al.*, Waggoner (1997) proposes using a cubic spline to approximate the forward rate function, with the number of nodes again being approximately one-third of the number of bonds in the sample, and spaced so that there is an equal number of bonds maturing between adjacent nodes. The Waggoner approach is termed the 'variable roughness penalty method' (VRP). The cubic spline forward rate curve is selected that will minimise the function (6.19):

$$\sum_{i=1}^{N}\left(P_i - \hat{P}_i(f)\right)^2 + \int_0^{\tau_N} \lambda(s)(f''(s))^2 \mathrm{d}s. \tag{6.19}$$

The roughness penalty λ is set as follows:

$$\lambda(t) = \begin{cases} 0.1 & 0 \le t \le 1 \\ 100 & 1 \le t \le 10 \\ 100{,}000 & 10 \le t \end{cases}$$

where t is measured in years. The VRP method is non-linear and can be solved using the non-linear least squares method.

6.4.2 *The Anderson–Sleath model*

In this section we summarise a paper by Anderson and Sleath which first appeared in the Bank of England *Quarterly Bulletin* in November 1999. The main objective of this work was to evaluate the relative efficacy of parametric versus spline-based methods. In fact different applications call for different methods; the main advantage of spline methods is that individual functions in between knot points may move in fairly independent fashion, which makes the resulting curve more flexible than that possible using parametric techniques. In Section 6.5.1 we reproduce their results with permission, which shows that a shock introduced at one end of the curve produces unsatisfactory results in the parametric curve.

The Anderson–Sleath model, which is the method adopted by the Bank of England, is a modification of the Waggoner approach in a number of significant ways. The $\lambda(t)$ function of Waggoner was adapted thus:

$$log\ \lambda(m) = L - (L - S)e^{-m/\mu} \tag{6.20}$$

where the parameters to be estimated are L, S, μ. In addition the difference in bond market and theoretical prices is weighted with the inverse of the modified duration of the bond. This accounts for observed pricing errors for bonds that are more volatile than others.

The model therefore minimises the expression in (6.21):

$$X = \sum_{i=1}^{N}\left(\frac{P_i - \pi_i(c)}{MD_i}\right)^2 + \int_0^{M} \lambda_t(m)(f''(m))^2 \mathrm{d}m \tag{6.21}$$

where P and MD are the price and modified duration of bond i, c is the parameter vector of the polynomial spline being estimated and M is the time to maturity of the longest-dated bond.

The outstanding feature of the Anderson–Sleath approach is their adaptation of both spline and parametric techniques.

6.4.3 *Applications*

Each of the methods described in this section can be used to fit the zero-coupon curve with validity. In practice results produced by each method imply that certain techniques are more suitable than others under specific conditions. Generally the incorporation of a 'roughness' penalty that varies across maturities produces more accurate pricing of short-dated bonds, and this is the case in the Fisher *et al.* and Waggoner methods. The McCulloch technique is reasonably accurate and, as it is a linear method, is more straightforward to implement than the other techniques. It produces a similar curve to the VRP method in terms of goodness of fit and smoothness. Therefore in most cases it is reasonable to use this method. The advantage of the VRP method is that it allows the user to select the degree of smoothing.

In deciding which method to use, practitioners will need to consider the effectiveness of each approach with regard to flexibility, simplicity and consistency. The requirements of central monetary authorities differ in some respects to investment and commercial banks, as we noted at the start of the chapter. Generally however curves should fit as wide a range of term structures as possible, and be tractable, or straightforward to compute. They should also be consistent with a yield curve model. For example the approach presented by Bjork and Christensen (1997) is compatible with the Hull–White or extended Vasicek yield curve model. In the same paper it is stated that the Nelson and Siegel technique is not consistent with any common term structure model. James and Webber (2000) state that the simplicity of the Nelson and Siegel approach, which is an advantage of the technique, is also its main drawback. In the same review it is concluded that B-spline methods are the most flexible and consistent, along with that described by Bjork and Christensen.

6.5 The Anderson–Sleath evaluation

6.5.1 *Fitting the spot curve*

In this section we summarise, with permission, results obtained in highly innovative research by Anderson and Sleath (1999), comparing the different methods. The accuracy of any of the techniques is usually tested by using a goodness of fit measure, for example if we fit the curve using n bonds we wish to minimise the measure given by (6.22):

$$X_P = \sum_{i=1}^{N} \left(\frac{P_i - \prod_i(\rho)}{MD_i} \right)^2 \tag{6.22}$$

where

P_i is the market price of the ith bond
MD_i is the modified duration of the ith bond
$\prod_i(\rho)$ is the fitted price of the ith bond.

A popular technique is the spline-based method of curve fitting. Unlike other methods (such as the parametric Svensson method) which specify a single short-rate to describe the instantaneous forward rate curve, spline-based methods fit a curve to observed data that is composed of a number of sections, but with constraints to ensure that the curve is smooth and continuous. As this is one of the aims we stated at the beginning, this is an advantage of the spline-based method, as it allows individual sections of the curve to move independently of each other. This is demonstrated in Figures 6.3 and 6.4, which show a hypothetical yield curve that has been fitted, from an assumed set of bond prices, using the cubic spline method and a parametric method such as Svensson. The change of the long bond yield has a significant effect on the Svensson curve, notably at the short end of the curve. The spline curve however undergoes only a slight change in response to the change in yield, and only at the long end.

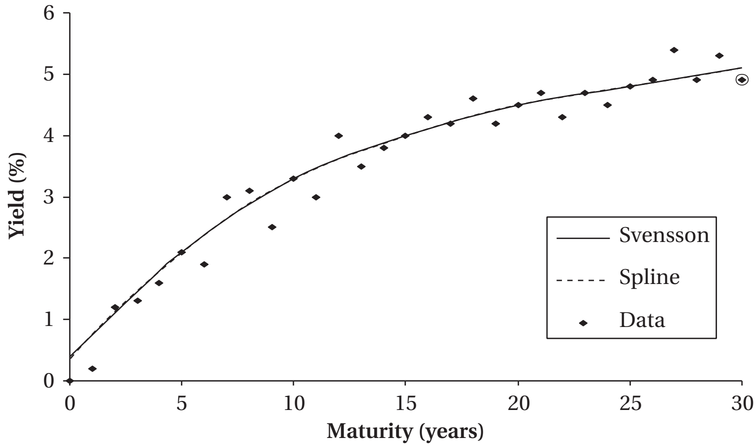


Figure 6.3: Yield curves fitted using cubic spline method and Svensson parametric method, hypothetical bond yields. Reproduced with permission from the Bank of England *Quarterly Bulletin*, November 1999.

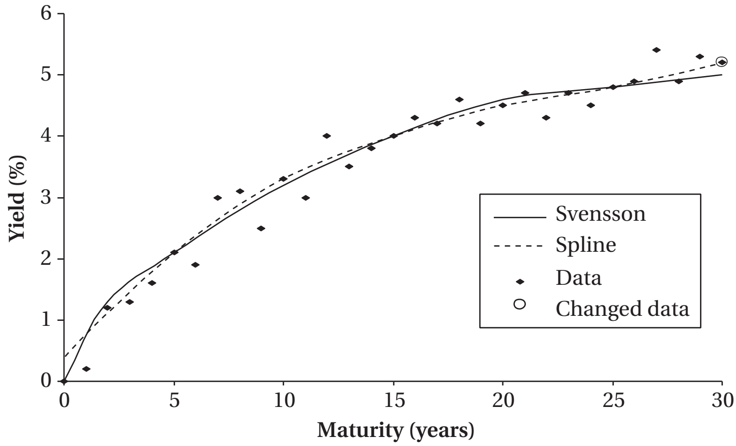


Figure 6.4: Effect on fitted yield curves of change in long-dated bond yield. Reproduced with permission from the Bank of England *Quarterly Bulletin*, November 1999.

The effect of a change in yield on the Svensson curve is amplified because the technique specifies a constraint that results in yields converging to a constant level. This assumption is based on the belief that forward rates reflect market expectations of the future level of short rates, and following this the 30-year forward rate will be expected to be not significantly different from the 25-year or 20-year forward rate. This causes the forward rate after about 10 years to converge to a constant level.

We can compare fitted yield curves to an actual spot rate curve wherever there is an active government (risk-free) zero-coupon market in operation. In the UK a zero-coupon bond market was introduced in December 1997. In theory any derived spot rate curve can be compared to the actual spot rate curve, this comparison serving to provide an instant check of the accuracy of the yield curve model. In practice however, discrepancies between the observed and fitted curves may not have that much significance, because of the way that strip yields behave in practice. In the UK market there is a certain level of illiquidity associated with strip yields at certain points of the term structure; the UK market also exhibits a common trait of strip markets everywhere that the longest-dated issue traded dear to the yield curve. Another factor is that coupon strips trade cheaper to principal strips; which yield should be used in the comparison?[6] In Figure 6.5 we compare the theoretical spot curve fitted using the Svensson method to the observed coupon strip curve in July 1998, at a time when the UK gilt yield curve was inverted.

The fitted curve exhibits the constant long yield that we observed in the hypothetical yield curve in Figure 6.1, while the strip curve trades expensive at the long end, which as we noted is a common observation. Nevertheless for the purposes of accurate fitting the parametric method exhibits a significant difference to the observed curve. A cubic spline-based fitted curve such as that proposed by Waggoner (1997) produces a more realistic curve, as shown in Figure 6.6.

This reflects the properties of the spline curve, including the fact that forward rates are described by a series of segments that are in effect connected together. This has the effect of localising the influence of individual yield movements to only the relevant part of the yield

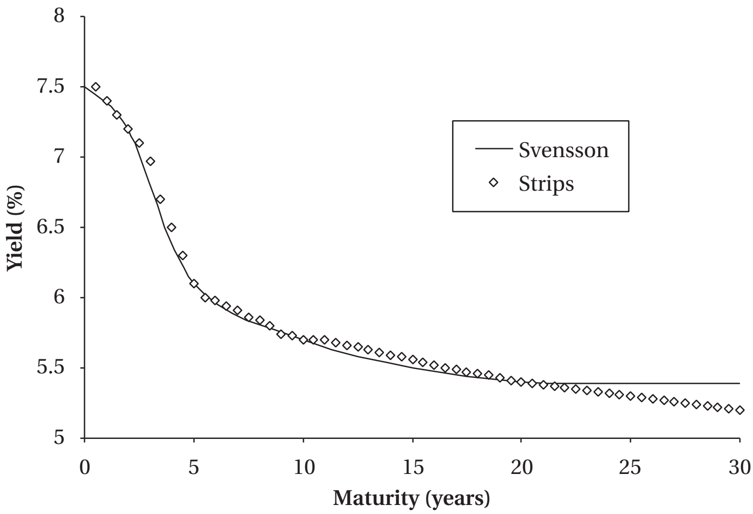


Figure 6.5: Comparison of fitted spot yield curve to observed spot yield curve. Reproduced with permission from the Bank of England *Quarterly Bulletin*, November 1999.

6 This is the observation that, due to demand and liquidity reasons, zero-coupon bonds sourced from the principal cash flow of a coupon bond trade at a lower yield than equivalent-maturity zero-coupon bonds sourced from the coupon cash flow of a conventional bond.

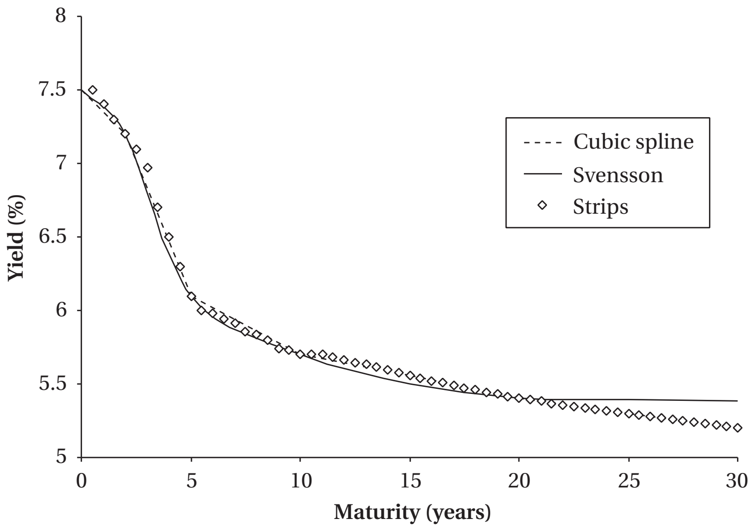


Figure 6.6: Fitted yield curves and observed strip yield curve, July 1998. Reproduced with permission from the Bank of England *Quarterly Bulletin*, November 1999.

curve; it also allows the curve to match more closely the observed yield curve. The goodness of the spline-based method is measured using (6.23):

$$X_s = X_P + \int_0^M \lambda_t(m)(f''(m))^2 \mathrm{d}m \tag{6.23}$$

where $f''(m)$ is the second derivative of the fitted forward curve and M is the maturity of the longest-dated bond. The term $\lambda_t(m)$ is the 'roughness penalty'. Figure 6.6 shows that the spline-based method generates a more realistic curve, that better mirrors the strip yield curve seen in Figure 6.5.

6.5.2 *Repo and estimating the short-end of the yield curve*

For the purposes of conducting monetary policy and for central government requirements, little use is made of the short-end of the yield curve. This is for two reasons; one is that monetary and government policy is primarily concerned with medium-term views, for which a short-term curve has no practical input, the second is that there is often a shortage of data that can be used to fit the short-term curve accurately. In the same way that the long-term term structure must be fitted using risk-free instruments, the short-term curve can only be estimated using Treasury bills. The T-bill can be restricted to only a small number of participants in some markets; moreover the yield available on T-bills reflects its near cash, risk-free status, and so may not be the ideal instrument to use when seeking to extract market views on forward rates. So for liquidity purposes the existence of an alternative instrument to T-bills would be useful. In most respects the government repurchase market or repo market is a satisfactory substitute for T-bills, although there is an element of counterparty risk associated with repo that does not apply to T-bills, they can be considered to be essentially risk-free instruments, more so if margin has been taken by the party

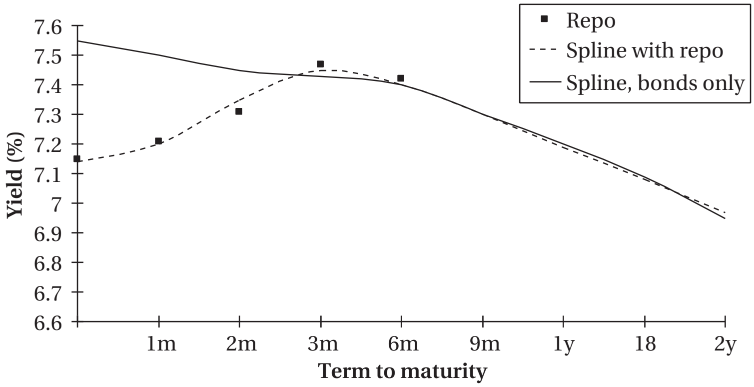


Figure 6.7: Fitting short-term yield curves using government repo rates. Reproduced with permission from the Bank of England *Quarterly Bulletin*, November 1999.

lending cash. We can therefore consider general collateral repo to be essentially a liquid, short-term and risk-free instrument.[7]

The fitted spot curve can differ considerably if yields on short-term repo are included. The effect is shown in Figure 6.7, which is reproduced from Anderson and Sleath (1999). Note that this is a short-term spot curve only; the maturity extends out to only two years. Two curves have been estimated; the cubic spline-based yield curve using the repo rate and without the repo rate. The curve that uses repo data generates a curve that is much closer to the money market yield curve than the one that does not. The only impact is at the very short end. After about one year, both approaches generate very similar curves.

For an account of the impact of 'special' repo rates on term structure modelling see Barone and Risa (1994) and Duffie (1993), which are available from the respective institution Web sites.

Appendices

Appendix 6.1: The McCulloch cubic spline model

This was first described by McCulloch (1975) and is referred to in Deacon and Derry (1994). We assume the maturity term structure is partitioned into q knot points with $q_l, \ldots, q_q$ where $q_l = 0$ and q_q is the maturity of the longest-dated bond. The remaining knot points are spaced such that there is, as far as possible, an equal number of bonds between each pair of knot points. With $j < q$, we employ the following functions:

- for $m < q_{j-1}$

$$f_j(m) = 0 \tag{6.24}$$

[7] See Choudhry (2002) for more information on the repo markets and the UK gilt repo market.

- for $q_{j-1} \leq m \leq q_j$

$$f_j(m) = \frac{(m - q_{j-1})^3}{6(q_j - q_{j-1})} \tag{6.25}$$

- for $q_i \leq m \leq q_{j+1}$

$$f_j(m) = \frac{c^2}{6} + \frac{cf}{2} + \frac{f^2}{2} - \frac{f^3}{6(q_{j+1} - q_j)} \tag{6.26}$$

- where

$$c = q_j - q_{j-1}$$
$$f = m - q_j$$

- for $q_{j+1} \leq m$

$$f_j(m) = (q_{j+1} - q_{j-1})\left(\frac{2q_{j+1} - q_j - q_{j-1}}{6} + \frac{m - q_{j+1}}{2}\right) \tag{6.27}$$

- for $j = q$

the function $f_q(m) = m$ for all values of m.

Appendix 6.2: Parametric and cubic spline yield curve models

In the Nelson and Siegel method (1987), we may model the implied forward rate yield curve along the entire term structure using the following function:

$$rf(m,\ \beta) = \beta_0 + \beta_1\ exp\left(\frac{-m}{t_1}\right) + \beta_2\left(\frac{m}{t_1}\right)exp\left(\frac{-m}{t_1}\right) \tag{6.28}$$

where $\beta = (\beta_0,\ \beta_1,\ \beta_2,\ t_1)'$ is the vector of parameters describing the yield curve, and m is the maturity at which the forward rate is calculated. There are three components, the constant term, a decay term and a term reflecting the 'humped' nature of the curve. The shape of the curve will gradually lead into an asymptote at the long end, the value of which is given by β_0, with a value of $\beta_0 + \beta_1$ at the short end.

Svensson (1994) presents a modification of this, by means of an adjustment to allow for the humped characteristic of most yield curves. This is fitted by adding an extension, as shown by (6.29):

$$rf(m,\ \beta) = \beta_0 + \beta_1\ exp\left(\frac{-m}{t_1}\right) + \beta_2\left(\frac{m}{t_1}\right)exp\left(\frac{-m}{t_1}\right) + \beta_3\left(\frac{m}{t_2}\right)exp\left(\frac{-m}{t_2}\right). \tag{6.29}$$

So we note that the Svensson curve is modelled using six parameters, the additional inputs being β_3 and t_2.

A different approach is adopted by smoothing cubic spline models. A generic spline is a segmented polynomial, or a curve that is constructed from individual polynomial segments that are joined together at user-specified 'knot points'. That is, the x-axis is divided into selected segments (the knot points). The segments can be at equal intervals or otherwise. At the knot points the curve and its first derivative are continuous at all points along the curve. Generally the market uses cubic functions, resulting in a cubic spline. A cubic spline is given by (6.30):

$$S(x) = ax^3 + \beta x^2 + \gamma x + \delta + \sum_{i=1}^{N-1} \eta_i |x - k_i|^3 \tag{6.30}$$

for a range of constants α, β, γ, δ, η and where k_i, $i = [0, N]$ is the set of knot points. The expression in (6.30) is the most common one used for a cubic spline, however in practice it is unwieldy for the purposes of calculation. Therefore splines are usually constructed as a linear combination of cubic basis splines or B-splines. This is a general transformation which removes the numerical problems associated with (6.30). A B-spline of order n can be written in the form:

$$B_{i,\,n}(x) = \frac{x - k_i}{k_{i+n-1} - k_i} B_{i,\,n-1}(x) + \frac{k_{i+n} - x}{k_{i+n} - k_{i+1}} B_{i+1,\,n-1}(x) \tag{6.31}$$

where $B_{i,\,1}(x) = 1$ if $k_i \leq x < k_{i+1}$, and $B_{i,\,1}(x) = 0$ otherwise. This approach was described in Lancaster and Šalkauskas (1986). When a large number of knot points are used a cubic spline can be used for interpolation, however as noted by Anderson and Sleath (1999) this approach is not used for monetary policy purposes, because it does not produce a smooth curve.

Selected bibliography and references

Anderson, N., Breedon, F., Deacon, M., Derry, A., Murphy, M., *Estimating and Interpreting the Yield Curve*, Wiley, 1996.

Anderson, N., Sleath, J., 'New estimates of the UK real and nominal yield curves', *Bank of England Quarterly Bulletin*, November 1999, pp. 384–392.

Bank of England, *Quarterly Bulletin*, August 1999.

Barone, E., Risa, S., 'Valuation of floaters and options on floaters under special repo rates', Instituto Mobiliare Italiano, 1994.

Bjork, T., Christensen, B., 'Interest rate dynamics and consistent forward rate curves', University of Aarhus Working Paper, 1997.

Bliss, R., 'Testing term structure estimation methods', *Advances in Futures and Options Research* 9, 1997, pp. 197–231.

Brenner, R., Jarrow, R., 'A simple formula for options on discount bonds', *Advances in Futures and Options Research* 6, 1993, pp. 45–51.

Choudhry, M., *Bond Market Securities*, FT Prentice Hall, 2001.

Choudhry, M., *The Repo Handbook*, Butterworth-Heinemann, 2002.

Cox, J., Ingersoll, J., Ross, S., 'A theory of the term structure of interest rates', *Econometrica* 53, 1985, pp. 385–407.

Dahlquist, M., Svensson, L., 'Estimating the term structure of interest rates for monetary policy analysis', *Scandinavian Journal of Economics* 98, 1996, pp. 163–183.

Deacon, M., Derry, A., 'Estimating the term structure of interest rates', *Bank of England Working Paper*, No. 24, July 1994.

Dothan, U., 'On the term structure of interest rates', *Journal of Financial Economics* 6, 1978, pp. 59–69.

Duffie, D., *Special Repo Rates*, Graduate School of Business, Stanford University, 1993.

Edmister, R.O., Madan, D.P., 'Informational content in interest rate term structures', *The Review of Economics and Statistics* 75(4), November 1993, pp. 695–699.

Fisher, M., Nychka, D., Zervos, D., 'Fitting the term structure of interest rates with smoothing splines', *Finance and Economic Discussion Series* 95–1, Federal Reserve Board, 1995.

Harrison, J., Pliska, S., 'Martingales and stochastic integrals in the theory of continuous trading', *Stochastic Processes and their Applications* 11, 1981, pp. 215–260.

Heath, D., Jarrow, R., Morton, A., 'A new methodology for contingent claims valuation', *Econometrica* 60, 1992, pp. 76–106.

James, J., Webber, N., *Interest Rate Modelling*, Wiley, 2000.

Kim, J., 'A discrete-time approximation of a one-factor Markov model of the term structure of interest rates', Graduate School of Industrial Administration, Carnegie-Mellon University, 1993.

Lancaster, P., Šalkauskas, K., *Curve and Surface Fitting: An Introduction*, Academic Press, 1986.

Mastronikola, K., 'Yield curves for gilt-edged stocks: a new model', *Bank of England Discussion Paper* (*Technical Series*), No. 49, 1991.

McCulloch, J., 'Measuring the term structure of interest rates', *Journal of Business* 44, January 1971, pp. 19–31.

McCulloch, J., 'The tax adjusted yield curve', *Journal of Finance* 30, June 1975, pp. 811–829.

Nelson, C., Siegel, A., 'Parsimonious modelling of yield curves', *Journal of Business* 60, 1987, p. 473.

Schaefer, S., 'Measuring a tax-specific term structure of interest rates in the market for British government securities', *The Economic Journal* 91, June 1981, pp. 415–438.

Shea, G., 'Interest-rate term structure estimation with exponential splines: a note', *Journal of Finance* 40(1), March 1985, pp. 319–325.

Suits, D., Mason, A, Chan, L., 'Spline functions fitted by standard regression methods', *Review of Economics and Statistics* 60, 1978, pp. 132–139.

Svensson, L., 'Estimating and interpreting forward interest rates: Sweden 1992–94', International Monetary Fund *Working Paper* 114, 1994.

Svensson, L., 'Estimating forward interest rates with the extended Nelson & Siegel method', *Sveriges Riksbank Quarterly Review* 3, 1995, p. 13.

Vasicek, O., 'An equilibrium characterization of the term structure', *Journal of Financial Economics* 5, 1977, pp. 177–188.

Vasicek, O., Fong, H.G., 'Term structure modelling using exponential splines', *Journal of Finance* 37(2), May 1982, pp. 339–356.

Waggoner, D., 'The robustness of yield and forward rate extraction methods', Unpublished Working Paper, Federal Reserve Bank of Atlanta, September 1996.

Waggoner, D., 'Spline methods for extracting interest rate curves from coupon bond prices', *Working Paper*, No. 97–10, Federal Reserve Bank of Atlanta, 1997.

Wiseman, J., 'The exponential yield curve model', European Fixed Income Research, Technical Specification, JP Morgan, 1994.

Zheng, C., 'An arbitrage-free SAINTS model of interest rates', *Working Paper*, First National Bank of Chicago, 1994.

7 Estimating and Interpreting the Term Structure II: a Practical Implementation of the Cubic Spline Method[1]

We noted in Chapter 4 that the term structure of interest rates defines the set of spot or zero-coupon rates that exist in a debt capital market, of default-free bonds, distinguished only by their term to maturity. In practice the term structure is defined as the array of discount factors on the same maturity term. Extracting the term structure from market interest rates has been the focus of extensive research, reflecting its importance in the field of finance. We also noted that the term structure is used by market practitioners for valuation purposes and by central banks for forecasting purposes. The accurate fitting of the term structure is vital to the smooth functioning of the market. Of the methods proposed for fitting the term structure, practitioners desire an approach that is accessible, straightforward to implement and as accurate as possible. In general there are two classes of curve fitting techniques; the *parametric* methods, so-called because they attempt to model the yield curve using a parametric function; and the *spline* methods.[2] Parametric methods include the Nelson–Siegel model and a modification of this proposed by Svensson (1994, 1995), as well as models described by Wiseman (1994) and Bjork and Christensen (1997). James and Webber (2000) suggest that these methods produce a satisfactory overall shape for the term structure but are suitable only where good accuracy is not required. Market practitioners instead generally prefer an approach that gives a reasonable trade-off between accuracy and ease of implementation, an issue we explore in this chapter.

The cubic spline process presents no conceptual problems, and is an approximation of the market discount function. McCulloch (1975) uses cubic splines and Beim (1992) states that this approach performs at least as satisfactorily as other methods. Although the basic approach can lead to unrealistic shapes for the forward curve (for example, see Vasicek and Fong (1982) and their suggested improvement on the approach using exponential splines),

[1] This chapter was co-authored with Rod Pienaar of Deutsche Bank AG, London. The views and opinions expressed remain those of the authors in their individual private capacity, and do not represent the views of any employing institution. It was published previously in Fabozzi, F., (ed.), *Interest Rate, Term Structure and Valuation Modeling*, John Wiley & Sons, 2002.

[2] Parametric models are also known as *parsimonious* models.

it is an accessible method and one that gives reasonable accuracy for the spot rate curve. Adams and van Deventer (1994) illustrate using the technique to obtain maximum smoothness for forward curves (and an extension to *quartic* splines), while the basic technique has been improved as described by Fisher, Nychka and Zervos (1995), Waggoner (1997) and Anderson and Sleath (1999). These references are considered later.

Splines are a non-parametric polynomial interpolation method.[3] There is more than one way of fitting them. The simplest method is an ordinary least squares regression spline, but this approach produces wildly oscillating curves. The more satisfactory is a smoothing splines method. Following on from our overview of fitting techniques in Chapter 6, in this chapter we consider the cubic spline approach and how to implement it.

7.1 Cubic splines

7.1.1 *Fitting a discount function*

In mathematics a spline is a piecewise polynomial function, made up of individual polynomial sections or segments that are joined together at (user-selected) points known as *knot points*. Splines used in term structure modelling are generally made up with cubic polynomials, and the reason for cubic polynomials, as opposed to polynomials of order say, two or five, is explained in straightforward fashion by de la Grandville (2001). A cubic spline is a function of order three, and a piecewise cubic polynomial that is twice differentiable at each knot point. At each knot point the slope and curvature of the curve on either side must match. We employ the cubic spline approach to fit a smooth curve to bond prices (yields) given by the term discount factors.

A polynomial of sufficiently high order may be used to approximate to varying degrees of accuracy any continuous function, which is why a polynomial approximation of a yield curve may be attempted. For example James and Webber (2000) state that given a set of m points with distinct values, a Lagrange polynomial of degree m will pass through every point.[4] However, the fit can be very wild with extreme behaviour at the long end. We demonstrate how a cubic spline approximation can be used to obtain better results.

This chapter provides a discussion of piecewise cubic spline interpolation methodology and its application to the term structure. We recommend a cubic spline technique because this ensures that the curve passes through all the selected (market determined) node points. This enables practitioners to fit a yield curve to observed market rates (Libor or bond yields) reasonably accurately and produces a satisfactory zero coupon curve under most circumstances.

Our starting point is a set of zero curve tenors (or discount factors) obtained from a collection of market instruments such as cash deposits, futures, swaps or coupon bonds. We therefore have a set of tenor points and their respective zero rates (or discount factors). The mathematics of cubic splines is straightforward but we assume a basic understanding of calculus and a familiarity with solving simultaneous linear equations by substitution. An account of the methods analysed in this chapter is given in Burden and Faires (1997), which has very accessible text on cubic spline interpolation.

3 A spline originally referred to a tool used by draughtsmen or carpenters for drawing smooth curves.

4 James and Webber (2000), pp. 430–432.

7.1.2 *Background on cubic splines*

When fitting a curve by interpolating between nodes or tenor points, the user must consider conflicting issues. There is a need to balance between simplicity and correctness, and hence a trade off between ease of use and the accuracy of the result. In certain cases practitioners will accept a lower degree of accuracy at the nodes, in favour of smoothness across the curve. In the cubic spline approach the primary aim is smoothness. In an attempt to create a smooth and accurate measurement at the nodes, however, we may be confronted by oscillation in the curve. Although linear interpolation is a reasonable calculation method, interest rate markets are not linear environments made up of coupled straight lines. The point between two tenors cannot be accurately estimated using a straight line.

Although there are a number of alternative methods available to the practitioner, a reasonable approach is to stick with the concept of piecewise interpolation but to abandon the use of straight lines. The reason that we do not depart from piecewise interpolation, is that this method of curve smoothing provides accuracy at the nodes because each piecewise function touches a node. Accuracy at the nodes can be an important consideration when a pricing methodology based on the elimination of arbitrage is employed. Thus we continue with piecewise fitting, but instead of applying a linear fitting technique, we apply a cubic polynomial to each piece of the interpolation. Cubic splines provide a great deal of flexibility in creating a continuous smooth curve both between and at tenor points.[5]

7.2 Cubic spline methodology

We assume that the practitioner has already calculated a set of nodes using a yield curve construction technique such as bootstrapping. A zero curve is then fitted using the cubic spline methodology by interpolating between nodes using individual cubic polynomials. Each polynomial has its own parameters but are constructed in such a way that their ends touch each node at the start and end of the polynomial. The set of splines, which touch at the nodes, therefore form a continuous curve. Our objective is to produce a continuous curve, joining market observed rates as smoothly as possible, which is the most straightforward means by which we can deduce meaningful data on the correct interest rate term structure in the market.

In Figure 7.1 we can see that two cubic polynomials which join at point x_{N+1} are used to form a continuous curve. However, it is also clear from the curves in Figure 7.1 that the two polynomials do not result in a smooth curve. In order to have a smooth curve we need to establish 'smoothing' criteria for each spline. To do this we must first ensure that the polynomials touch or join together at the nodes. Secondly, we must ensure that where the polynomials touch the curve is smooth. Finally, we ensure that the curve is continuously differentiable or, in other words, the curve has a smooth rate of change at and between tenor points. The required criteria to meet these conditions are:

Requirement 1: the value of each polynomial is equal at tenor points;
Requirement 2: the first differential of each polynomial is equal at tenor points;
Requirement 3: the second differential of each polynomial is equal at tenor points; and
Requirement 4: the second differential of each polynomial is continuous between tenor points.

[5] See the earlier footnote for a word on the origin of the use of the term 'spline'.

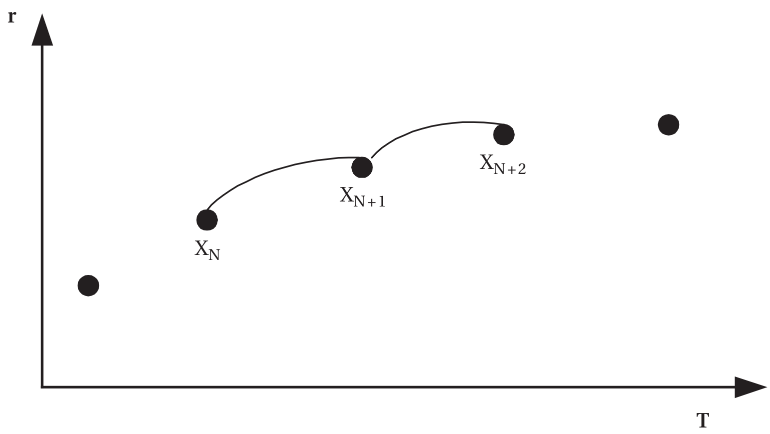


Figure 7.1: Forming a continuous curve, cubic polynomials.

Considering a polynomial of the form $y = ax^3 + bx^2 + cx + d$, the second differential $y'' = 6ax + 2b$ is a linear function and by its very definition is continuous between tenor points. The fourth requirement is therefore always met and this chapter will not deal with this requirement in any further detail. The rest of this chapter will refer to the first three requirements and how they are met at the nodes.

7.3 The supposition

Assuming the final solution is unknown at this stage, it seems plausible that an almost infinite set of parameters *a, b* and *c* can be found which will result in all of our cubic spline requirements being met.

We observe in Figure 7.2 three imaginary curves, all of which would meet our requirements that the:

- first differential of each spline is equal at tenor points; and
- second differential of each spline is equal at tenor points.

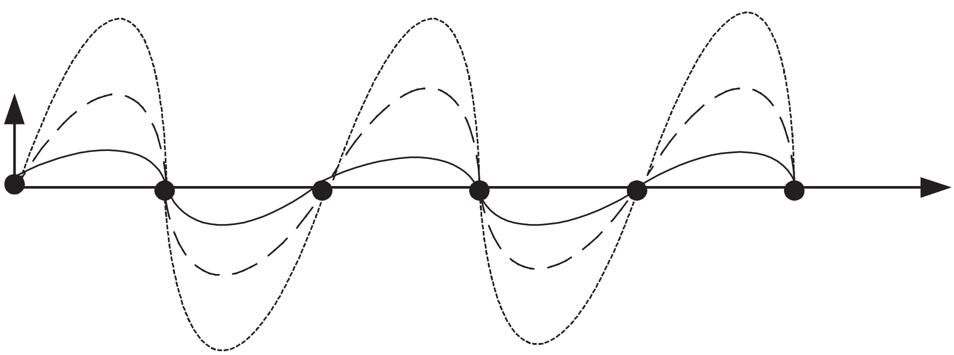


Figure 7.2: Three different possible solutions.

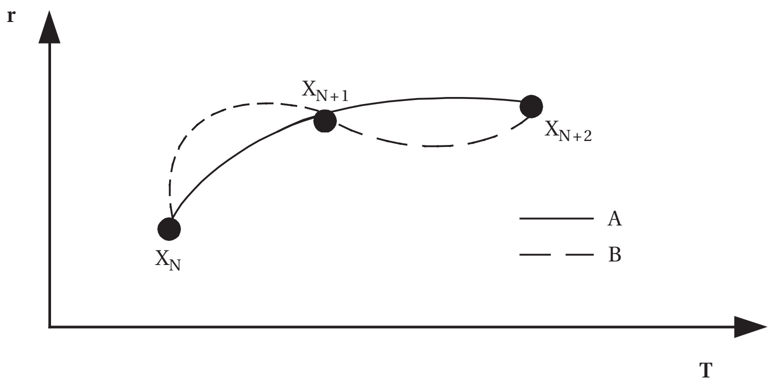


Figure 7.3: Two acceptable solutions.

Admittedly we have considered nodes that are sitting in a straight line but even where the nodes do not line up it may be possible to find a range of possible solutions. Taking this further, spline A and spline B as shown in Figure 7.3 are valid solutions yet it is intuitive, given our knowledge of interest rate markets, that A is likely to be more suitable for our purposes of yield curve interpolation.

The issue to determine therefore, is, is there an infinite set of parameters each of which would meet our requirements for fitting the curve; or is it possible to determine a single solution? Of course our requirement is in a single solution; moreover, a solution that can be found quickly from any set of market rates.

7.4 Practical approach

7.4.1 *A working environment*

By splitting the yield curve into individual node/tenor pairs, we may work with individual lines within each tenor. A cubic polynomial can then be added to each line to provide the cubic spline. For ease of illustration, we take this one step further and imagine an alternative horizontal axis. This is referred to as ‘capital’ X as shown in Figure 7.4. Assume that between each node pair this horizontal axis X runs from *0* (at x_N) to $x_{N+1} - x_N$ (at x_{N+1}).

In Figure 7.4 the X axis is a calculated value determined from the x axis. The points x_N and x_{N+1} are isolated for spline S_N. It is then assumed that X_0 equals zero at x_N and stretches to X_N which equals $(x_{N+1} - x_N)$ on the X axis. If these lines are fully isolated then a cubic polynomial, of the form $y = aX^3 + bX^2 + cX + d$, can be constructed to touch the points x_N and x_{N+1}.

7.4.2 *The first requirement*

In order for the polynomial to touch the nodes then a cubic polynomial must be constructed so that at point X_0 the polynomial provides a result that is equal to y_N. This is very

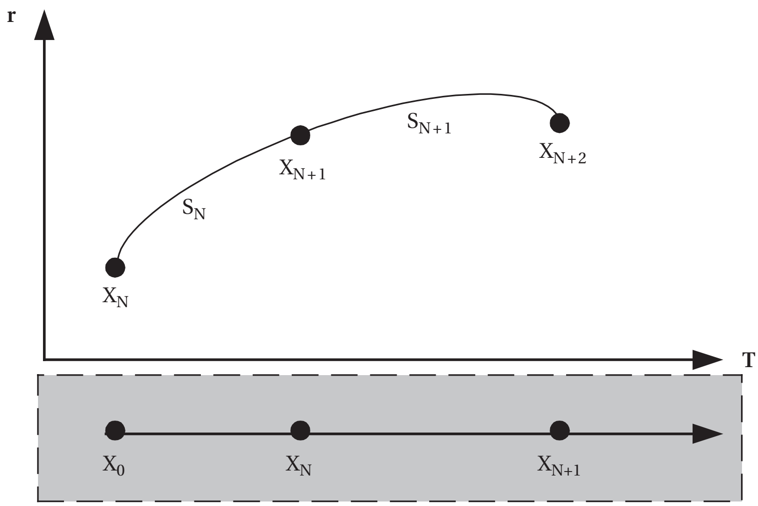


Figure 7.4: The imaginary 'X' axis.

easy to achieve. Since X is equal to zero at its starting point, the polynomial takes the following form:

$$y_N = a_N 0^3 + b_N 0^2 + c_N 0 + d_N$$

$$y_N = d_N$$

So as long as d_N is equal to y_N then our polynomial will touch the node at X_0.

In order for the polynomial to touch the second node, the node at point x_{N+1}, then the polynomial must take the following form at point X_N:

$$y_{N+1} = a_N(x_{N+1} - x_N)^3 + b_N(x_{N+1} - x_N)^2 + c_N(x_{N+1} - x_N) + d_N$$

OR

$$d_{N+1} = a_N X_N^3 + b_N X_N^2 + c_N X_N + d_N \tag{7.1}$$

where: $X_N = x_{N+1} - x_N$

It is worth noting that at this point in our process we do not know what the values of *a, b* or *c* are. These will be derived below from our other requirements.

7.4.3 *The second requirement*

To meet the second requirement of a cubic spline, the first differential y_N' must equal the first differential y_{N+1}' at the tenor point x_{N+1}.

In other words at node x_{N+1}:

$$3a_N X_N^2 + 2b_N X_N + c_N = 3a_{N+1} X_{N+1}^2 + 2b_{N+1} X_{N+1} + c_{N+1} \tag{7.2}$$

We know from our conditional working environment that at node x_{N+1} for function y_N' that $X = (x_{N+1} - x_N)$. We also know from the same assumption that $X = 0$ at the start of the next polynomial, i.e. for function y_{N+1}'. Therefore:

$$3a_{N+1}0^2 + 2b_{N+1}0 + c_{N+1} = 3a_N X_N^2 + 2b_N X_N + c_N$$

so that

$$c_{N+1} = 3a_N X_N^2 + 2b_N X_N + c_N \tag{7.3}$$

7.4.4 *Third requirement*

To meet the third requirement of a cubic spline, the second differential y_N'' assessed at the point x_{N+1} should equal the second differential y_{N+1}''.

In other words at node x_{N+1}:

$$6a_N X_N + 2b_N = 6a_{N+1}X_{N+1} + 2b_{N+1}$$

We know from our conditions that at node x_{N+1} for function y_N'' that $X = (x_{N+1} - x_N)$. We also know from the same assumption that $X = 0$ for function y_{N+1}''. Therefore:

$$6a_N X_N + 2b_N = 6a_{N+1}0 + 2b_{N+1}$$

$$6a_N X_N = 2b_{N+1} - 2b_N$$

$$a_N = \frac{b_{N+1} - b_N}{3X_N} \tag{7.4}$$

7.4.5 *Meeting all requirements simultaneously*

We now have equations which ensure that each of the requirements can be met. We now need a solution that will ensure that all requirements are met at the same time. By substitution a set of calculations can be performed which meet both requirements and reduce these equations down to a factor of parameter b only.

Using equation (7.4) as a substitute for a in equation (7.3) we obtain:

$$c_{N+1} = 3a_N X_N^2 + 2b_N X_N + c_N$$

$$c_{N+1} = \frac{3(b_{N+1} - b_N)}{3X_N}X_N^2 + 2b_N X_N + c_N$$

$$c_{N+1} = (b_{N+1} - b_N)X_N + 2b_N X_N + c_N$$

$$c_{N+1} = X_N(b_{N+1} + b_N) + c_N \tag{7.5}$$

Using equation (7.4) as a substitute for a in equation (7.1) we get:

$$d_{N+1} = \frac{(b_{N+1} - b_N)}{3X_N}X_N^3 + b_N X_N^2 + c_N X_N + d_N$$

$$d_{N+1} = \frac{(b_{N+1} - b_N)}{3}X_N^2 + b_N X_N^2 + c_N X_N + d_N$$

$$c_N X_N = -\frac{(b_{N+1} - b_N)}{3}X_N^2 - b_N X_N^2 + d_{N+1} - d_N$$

$$c_N = -X_N\frac{(b_{N+1} + 2b_N)}{3} + \frac{(d_{N+1} - d_N)}{X_N} \tag{7.6}$$

Taking this solution one step further we can substitute equation (7.6) into equation (7.5) as follows:

$$\frac{(d_{N+2}-d_{N+1})}{X_{N+1}}-X_{N+1}\frac{(b_{N+2}+2b_{N+1})}{3}=X_N(b_{N+1}+b_N)-X_N\frac{(b_{N+1}+2b_N)}{3}+\frac{(d_{N+1}-d_N)}{X_N}$$

$$-X_{N+1}(b_{N+2}+2b_{N+1})=3X_N(b_{N+1}+b_N)-X_N(b_{N+1}+2b_N)+3\frac{(d_{N+1}-d_N)}{X_N}-3\frac{(d_{N+2}-d_{N+1})}{X_{N+1}}$$

$$-X_{N+1}(b_{N+2}+2b_{N+1})=X_N(2b_{N+1}+b_N)+3\frac{(d_{N+1}-d_N)}{X_N}-3\frac{(d_{N+2}-d_{N+1})}{X_{N+1}}$$

$$X_{N+1}b_{N+2}=-X_N(2b_{N+1}+b_N)-3\frac{(d_{N+1}-d_N)}{X_N}+3\frac{(d_{N+2}-d_{N+1})}{X_{N+1}}-2X_{N+1}b_{N+1}$$

$$b_{N+2}=\frac{-2X_Nb_{N+1}-X_Nb_N-2X_{N+1}b_{N+1}-3\frac{(d_{N+1}-d_N)}{X_N}+3\frac{(d_{N+2}-d_{N+1})}{X_{N+1}}}{X_{N+1}} \tag{7.7}$$

7.4.6 *A unique solution*

For clarity and ease of illustration, the results of these equations are set out as a table of related formulas shown below in Table 7.1.

It is a simple matter to determine the values of parameters *a, b, c* and *d* at each node *n* by using the formulas set out in Table 7.1. Each node (from $n > 2$) is directly or indirectly dependent on the values of previous parameters and can be determined from those previous parameters. This is an important result, and means that any errors in the calculation early on are replicated and magnified throughout the analysis. However, the first two occurrences of *b* (b_0 and b_1) do not have previous nodes from which to determine their values. In other words, the only values for which we do not have solutions are those for b_0 and b_1.

X	**y (*d*)**	**Using equation 7.3 we can derive *a***	**Using equation 7.6 we can derive *b***	**Using equation 7.5 we can derive *c***
X_1	d_1	$\frac{b_2-b_1}{3X_1}$	$\mathbf{b_0}$	$-X_1\frac{(b_2+2b_1)}{3}+\frac{(d_2-d_1)}{X_1}$
X_2	d_2	$\frac{b_3-b_2}{3X_3}$	$\mathbf{b_1}$	$-X_2\frac{(b_3+2b_2)}{3}+\frac{(d_3-d_2)}{X_2}$
X_3	d_3	$\frac{b_4-b_3}{3X_4}$	$\frac{-2X_1b_2-X_1b_1-2X_2b_2-3\frac{(d_2-d_1)}{X_1}+3\frac{(d_3-d_2)}{X_2}}{X_2}$	$-X_3\frac{(b_4+2b_3)}{3}+\frac{(d_4-d_3)}{X_3}$
…	…	…	…	…
X_{N-1}	d_{N-1}	$\frac{b_N-b_{N-1}}{3X_{N-1}}$	$\frac{-2X_{N-3}b_{N-2}-X_{N-3}b_{N-3}-2X_{N-2}b_{N-2}-3\frac{(d_{N-2}-d_{N-3})}{X_{N-3}}+3\frac{(d_{N-1}-d_{N-2})}{X_{N-2}}}{X_{N-2}}$	$-X_{N-1}\frac{(b_N+2b_{N-1})}{3}+\frac{(d_N-d_{N-1})}{X_{N-1}}$
X_N	d_N	N/A	$\frac{-2X_{N-2}b_{N-1}-X_{N-2}b_{N-2}-2X_{N-1}b_{N-1}-3\frac{(d_{N-1}-d_{N-2})}{X_{N-2}}+3\frac{(d_N-d_{N-1})}{X_{N-1}}}{X_{N-1}}$	N/A

Table 7.1

Depending on the values assumed for b_0 and b_1, the result is *usually* an oscillating b and ever increasing $|b|$. This means that the slope of the spline gets steeper at each tenor as the absolute value of the first differential increases, so the slope of the curve oscillates.

This systematic wave, shown in Figure 7.5, is clearly not the kind of behaviour that is commonly observed in a yield curve and should therefore not be modelled into the curve. Furthermore, we have no unique solution at this stage. An infinite number of values can be assigned to b_0 and b_1 and therefore an infinite number of solutions can be obtained (most of which exhibit the depicted oscillation effect). So this is still not what we seek.

We need an additional restriction that allows us to find a single solution and which eliminates the oscillation of the output. The restriction that we put in place is to set the second differential of the first spline y_0'' and last spline y_N'' equal to a constant. We will use a constant of zero for now, but we come back to this constant at a later stage. Creating this additional restriction means that we are left with only one unknown, parameter b_2. This is demonstrated, using the constant zero, in Table 7.2.

If we find a value for b_2 that results in a final value of zero for b_N then we have a single solution and this solution should eliminate the oscillation shown above. We can determine this solution using two different methods:

- iteration; or
- Gaussian Elimination of a tri-diagonal matrix.

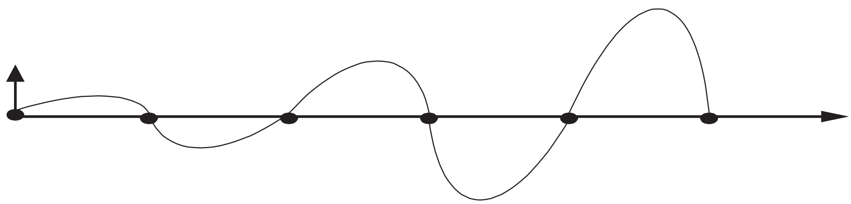


Figure 7.5: Increasing oscillation, and the systematic wave.

X	**y (d)**	**Using equation 7.3 we can derive a**	**Using equation 7.6 we can derive b**	**Using equation 7.5 we can derive c**
X_1	d_1	$\frac{b_2-b_1}{3X_1}$	**0**	$-X_1\frac{(b_2+2b_1)}{3}+\frac{(d_2-d_1)}{X_1}$
X_2	d_2	$\frac{b_3-b_2}{3X_3}$	***The only parameter left to solve for is b_1***	$-X_2\frac{(b_3+2b_2)}{3}+\frac{(d_3-d_2)}{X_2}$
X_3	d_3	$\frac{b_4-b_3}{3X_4}$	$\frac{-2X_1b_2-X_1b_1-2X_2b_2-3\frac{(d_2-d_1)}{X_1}+3\frac{(d_3-d_2)}{x_2}}{X_2}$	$-X_3\frac{(b_4+2b_3)}{3}+\frac{(d_4-d_3)}{X_3}$
...	...	...	...	...
X_{N-1}	d_{N-1}	$\frac{b_N-b_{N-1}}{3X_{N-1}}$	$\frac{-2X_{N-3}b_{N-2}-X_{N-3}b_{N-3}-2X_{N-2}b_{N-2}-3\frac{(d_{N-2}-d_{N-3})}{X_{N-3}}+3\frac{(d_{N-1}-d_{N-2})}{X_{N-2}}}{X_{N-2}}$	$-X_{N-1}\frac{(b_N+2b_{N-1})}{3}+\frac{(d_N-d_{N-1})}{X_{N-1}}$
X_N	d_N	N/A	**0**	N/A

Table 7.2

Before we consider each of these solution techniques we consider first the requirement of a boundary condition in order to obtain a unique solution for a cubic spline. In our discussion above we ordained a boundary condition of $b_0 = b_N = 0$. In practice two boundary conditions have become widely accepted:

1. **Natural spline:** In a natural spline the second differential at x_0 and x_N is set to zero. In other words $y_0'' = y_N'' = 0$.
2. **Clamped spline:** In a clamped spline the first differential of the function that produced the nodes and the first derivative of the spline are set equal. In other words $y_{0'} = f(x_0)'$ and $y_{N'} = f(x_N)'$. It is immediately apparent when we construct a yield curve that we do not have a function that can be used to replicate the nodes. The first differential of this function is therefore not available. A reasonable approximation can be used based on the slope of the linear interpolation function between the first two and the last two nodes. Although this provides a reasonable approximation in most circumstances it is not always an appropriate measure. An incorrect choice of boundary values could result in spurious and oscillating results at the short and/or long end of the curve.

An example using the same input data but different (albeit rather extreme) boundary values is shown in Figure 7.6. The natural boundary uses values zero and zero. In the clamped boundary we have used −50 and −50 as boundary values. Although these boundary values are extreme, they do illustrate the effect that inappropriate boundary values can have on spline results.

These results are not unexpected. Readers may question the practical difference between having a natural boundary condition against having a boundary condition that is obviously inappropriate. Both approaches may lead to oscillation and an incorrect result. The sole practical difference is that where we set our own boundary value, however inappropriate, the extent of the error is under our own control. For this reason users may prefer this approach.

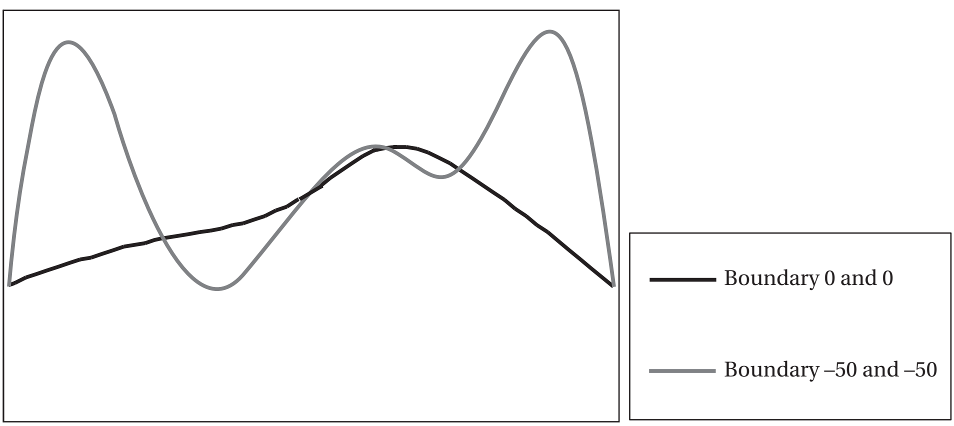


Figure 7.6: Natural and clamped splines.

7.4.7 *The solution*

We now consider each approach to obtaining the solution.

Iterative solution

A solution for b_1 can be obtained by iteration. This 'trial-and-error' style approach is straightforward to understand but is not without its limitations.

When a cubic spline solution is solved by iteration for a single parameter, the degree of accuracy required is very high. In test solutions the authors found that a higher degree of accuracy was required for a higher number of nodes. A calculation for fifteen nodes or more required the solution to be accurate to at least eight decimal places. Even apparently negligible differences in decimal accuracy can result in strange spline parameters and in turn produce the same oscillation observed above when no boundary values were set. This is particularly evident at the long end of the curve as the error becomes compounded by previous inaccuracies, thus leading to yield curves of limited practical application when anything longer than the medium-term maturity range is modelled.

A fictional set of numbers has been used to demonstrate this point in Table 7.3. The 'Date' column holds the maturity dates for each rate, while the 'Rate' column is of course the set of interest rates for each particular term to maturity.

This data is illustrated graphically in Figure 7.7.

In Table 7.3, an accuracy of eight decimal places is shown but in fact a much higher level (over 15 decimal places) of accuracy was required to calculate the results. When we adjust

Date	Rate (d)	parameter a	parameter b	parameter c
1-Jan-00	6.000	−0.00001228	0.00000000	0.00544212
7-Jan-00	6.030	0.00000351	−0.00022106	0.00411577
31-Jan-00	6.050	−0.00000019	0.00003181	−0.00042615
1-Apr-00	6.100	−0.00000001	−0.00000235	0.00137086
1-Jul-00	6.200	0.00000002	−0.00000426	0.00076898
1-Oct-00	6.250	−0.00000001	0.00000117	0.00048462
1-Jan-01	6.300	0.00000000	−0.00000042	0.00055340
1-Jul-01	6.400	−0.00000000	0.00000083	0.00062739
1-Jan-02	6.520	0.00000000	−0.00000126	0.00054853
1-Jan-03	6.610	−0.00000000	0.00000004	0.00010301
1-Jan-05	6.700	0.00000000	0.00000000	0.00013362
1-Jan-06	6.750	−0.00000000	0.00000003	0.00014328
1-Jan-07	6.800	0.00000000	−0.00000010	0.00011518
1-Jan-10	6.900	−0.00000000	0.00000014	0.00015545
1-Jan-11	6.960	0.00000000	−0.00000020	0.00013152
1-Jan-12	7.000	−0.00000000	0.00000023	0.00014041
1-Jan-14	7.100	0.00000000	−0.00000047	−0.00003778
1-Jan-15	7.050	−0.00000000	0.00000013	−0.00016286
1-Jan-20	7.000	0.00000000	−0.00000004	0.00000616
1-Jan-25	6.950	−0.00000000	0.00000002	−0.00002600
1-Jan-30	6.950		0.00000000	

Table 7.3

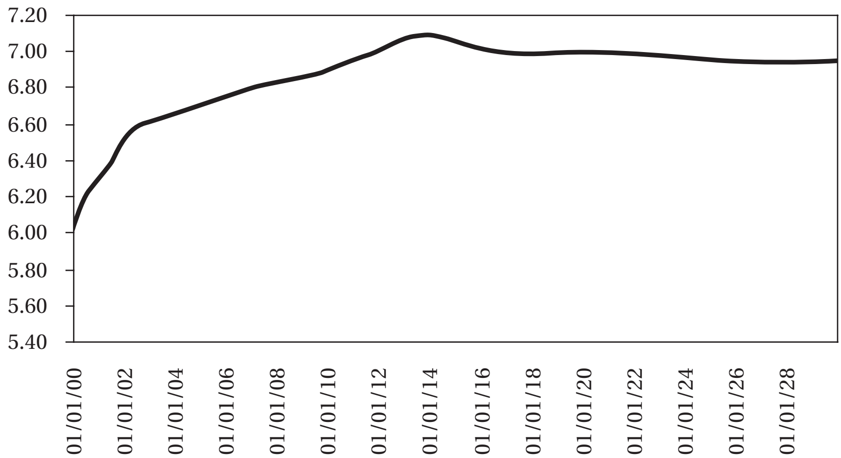


Figure 7.7: Graphic illustration of data in Table 7.3.

the level of accuracy, just on parameter b_1, to seven decimal places the results are significantly flawed, as shown in Table 7.4.[6]

It can be seen that within the long dates, parameter b starts to oscillate and grow in an exponential manner. A graphical representation of the rates as a result of this flawed data is shown in Figure 7.8. Note that the oscillation error is highly pronounced.

The degree of accuracy obtained through iteration is dependent on the starting point for the first calculation and the number of iterations allowed as a maximum. There is no way of ensuring that the required degree of accuracy will be obtained without undertaking very high magnitude (and process intensive) calculations in the iterative algorithm. Without the comfort of extensive manual review of the results by a person with a clear understanding of the calculation and its implications, we do not recommend the use of the iteration approach to derive a solution.

Solving for a system of linear equations by elimination

We now consider again equation (7.7) derived above, and re-arrange it slightly as (7.8).

$$X_{N+1}b_{N+2} + 2(X_N + X_{N+1})b_{N+1} + X_N b_N = -3\frac{(d_{N+1} - d_N)}{X_N} + 3\frac{(d_{N+2} - d_{N+1})}{X_{N+1}} \qquad (7.8)$$

It can be seen that all parameters X and d can be obtained by reference to values that are already known at the nodes. These are in fact node (or time-to-maturity) dependent constants. In other words we have a system of linear equations from node 1 to N. Readers will know that simultaneous linear equations can be solved by substitution. This method of solving linear equation can be applied to larger sets of linear equations, although we require increased processing power.

6 The results were calculated using the 'Goal Seek' function in Microsoft Excel.

Date	Rate (d)	parameter a	parameter b	parameter c
1-Jan-00	6.000	−0.00001228	0.00000000	0.00544210
7-Jan-00	6.030	0.00000351	−0.00022105	0.00411580
31-Jan-00	6.050	−0.00000019	0.00003179	−0.00042640
1-Apr-00	6.100	−0.00000001	−0.00000230	0.00137252
1-Jul-00	6.200	0.00000002	−0.00000442	0.00076105
1-Oct-00	6.250	−0.00000002	0.00000174	0.00051482
1-Jan-01	6.300	0.00000002	−0.00000255	0.00044055
1-Jul-01	6.400	−0.00000006	0.00000695	0.00123776
1-Jan-02	6.520	0.00000008	−0.00002345	−0.00179846
1-Jan-03	6.610	−0.00000011	0.00006372	0.01289764
1-Jan-05	6.700	0.00000103	−0.00017986	−0.07200383
1-Jan-06	6.750	−0.00000419	0.00095266	0.21006837
1-Jan-07	6.800	0.00000395	−0.00363079	−0.76744773
1-Jan-10	6.900	−0.00006704	0.00936251	5.51451411
1-Jan-11	6.960	0.00028391	−0.06404843	−14.44584982
1-Jan-12	7.000	−0.00043548	0.24683078	52.26970709
1-Jan-14	7.100	0.00407923	−0.70817417	−284.97230573
1-Jan-15	7.050	−0.00230683	3.75858533	828.42777079
1-Jan-20	7.000	0.00741195	−8.87822401	−8, 520.0324431
1-Jan-25	6.950	−0.02736125	31.74664902	33,260.580061
1-Jan-30	6.950		−118.13828171	

Table 7.4

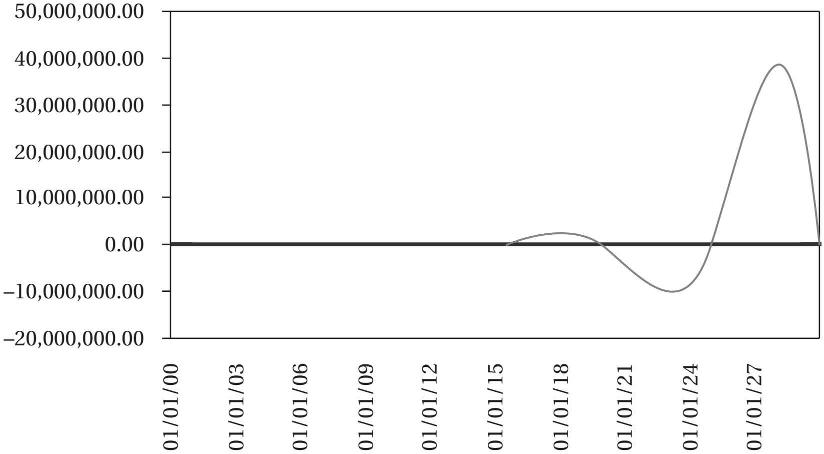


Figure 7.8: Graphic illustration of data in Table 7.4, showing excessive oscillation.

The system of equations can be represented in a $N-2$ by $N+1$ matrix as follows:

X_0	$2(X_0+X_1)$	X_1					$-3\left(\frac{(d_1-d_0)}{X_0}-\frac{(d_2-d_1)}{X_1}\right)$
	X_1	$2(X_1+X_2)$	X_2				$-3\left(\frac{(d_2-d_1)}{X_1}-\frac{(d_3-d_2)}{X_2}\right)$
		…	…	…			…
			…	…	…		…
				X_{N-2}	$2(X_{N-2}+X_{N-1})$	X_{N-1}	$-3\left(\frac{(d_{N-1}-d_{N-2})}{X_{N-2}}-\frac{(d_N-d_{N-1})}{X_{N-1}}\right)$

In essence, if you look at the parameters b for which we are attempting to solve, this can be laid over the above matrix as follows:

b_0	b_1	b_2				
	b_1	b_2	b_3			
		…	…	…		
			…	…	…	
				b_{N-2}	b_{N-1}	b_N

In other words, we are looking for a set of values for b_0 to b_N that will solve the linear system for each and every node N.

Our basic limitation imposed above is not lifted. We set b_0 and b_N equal to 0 in order to apply the natural boundary condition. We can then substitute our solution for equation/row 1 into equation/row 2. We perform a similar continuous set of substitutions until we have a solution for b_{N-1}. This solution can then be substituted backward through the solved equations to obtain a solution for b_1.

A matrix of this form, that is, an upper and lower triangular quadrant for which no value is required (observed by the grey shaded area) is also known as a *tri-diagonal matrix*. More advanced methods of solving matrices (and in particular tri-diagonal types) are available. It is outside the scope of this chapter to cover these methods in detail; interested readers may wish to consult Burden and Faires (1997). For the purposes of illustration, however, we have prepared a simple example solution for a small matrix of values, and this appears as an Appendix to this chapter.

The same values used for the iterative solution were processed using the elimination solution. The results and their illustrative chart are set out in Table 7.5 and Figure 7.9 respectively below.

On first observation these values appear to be identical to those obtained using the iterative solution. In fact, even at the highest level of accuracy possible in our iterative solution we notice a difference in the values for parameter **c** when we look at the dates 1 Jan 2014 onwards (which appear in the grey boxes in Table 7.5). Although this is not apparent in the chart, the results in the table where numbers appear with greater accuracy, show these and other small differences not shown in Figure 7.9.

Date	Rate (d)	parameter a	parameter b	parameter c
1-Jan-00	6.000	−0.00001228	0.00000000	0.00544212
7-Jan-00	6.030	0.00000351	−0.00022106	0.00411577
31-Jan-00	6.050	−0.00000019	0.00003181	−0.00042615
1-Apr-00	6.100	−0.00000001	−0.00000235	0.00137086
1-Jul-00	6.200	0.00000002	−0.00000426	0.00076898
1-Oct-00	6.250	−0.00000001	0.00000117	0.00048462
1-Jan-01	6.300	0.00000000	−0.00000042	0.00055340
1-Jul-01	6.400	−0.00000000	0.00000083	0.00062739
1-Jan-02	6.520	0.00000000	−0.00000126	0.00054853
1-Jan-03	6.610	−0.00000000	0.00000004	0.00010301
1-Jan-05	6.700	0.00000000	0.00000000	0.00013362
1-Jan-06	6.750	−0.00000000	0.00000003	0.00014328
1-Jan-07	6.800	0.00000000	−0.00000010	0.00011518
1-Jan-10	6.900	−0.00000000	0.00000014	0.00015545
1-Jan-11	6.960	0.00000000	−0.00000020	0.00013151
1-Jan-12	7.000	−0.00000000	0.00000023	0.00014041
1-Jan-14	7.100	0.00000000	−0.00000047	~~−0.00003779~~
1-Jan-15	7.050	−0.00000000	0.00000013	~~−0.00016284~~
1-Jan-20	7.000	0.00000000	−0.00000004	~~0.00000594~~
1-Jan-25	6.950	−0.00000000	0.00000002	~~−0.00002515~~
1-Jan-30	6.950		0.00000000	

Table 7.5

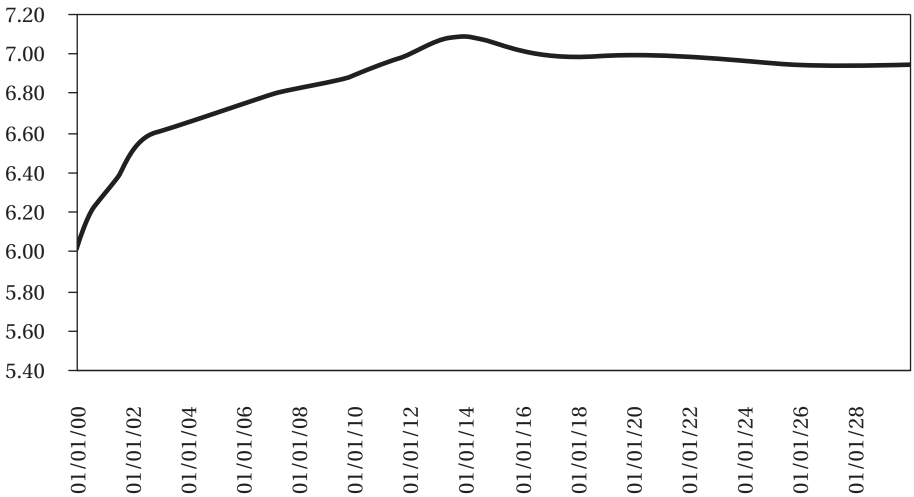


Figure 7.9: Graphic illustration of data in Table 7.5.

Based on these results we conclude that the technique of solving for a system of linear equations is superior to an iterative solution. This is because:

- no starting point for the calculation needs to be determined by the user or the system;
- the accuracy of the solution is not dependent on the number of iterative calculations performed; and
- the results do not need the same degree of review to assess their accuracy.

This is not to say that this method is flawless. Even a tri-diagonal methodology is reliant on the degree of precision applied in its calculation. Modern computing hardware and software has limitations in the size or length of the floating point numbers that it can process. However, if programmed with care, a typical application can deal with significantly large numbers.

7.5 Empirical proof of precision

The cubic spline application (CUBED[3]) is available on the website www.YieldCurve.com. This demonstrates the methodology we have described in this chapter. In this application we have chosen C++ as the programming language and we have used the C++ 'long double' variable type to store and process our values. A long double is usually anything between a 74 and 128 bit place holder, depending on the compiler and the system on which the calculations are performed. Applying some basic binary mathematics and allowing 1 bit for sign storage we can calculate:

$$2^{71} = 2,361,183,241,434,820,000,000$$

This should be sufficient to provide an adequate level of accuracy for most cubic spline calculations required of a zero curve application.[7] To test this we have performed empirical testing to corroborate our conclusion using a completely fictitious set of data that was designed to provide an extreme testing environment and data that is more sensitive to calculation anomalies than any likely to occur in real life.[8] Our fake input values were chosen to include:

- a large number of nodes (over 100);
- high oscillations at various points in the curve; and
- various points of flat data.

A large number of tenors was chosen to compound any rounding errors that might occur as part of the elimination multiplier. Oscillation at various points in the curve are used to set up waves that can continue when they subsequently flow into areas of flat data and which would highlight errors, if they occur. Flat sections of the curve are used so that any errors become highly visible.

---

7 This assurance is based on the fact that a typical yield curve application very, very rarely has more than 30 nodes. Any application where there are large node numbers may require higher levels of accuracy.

8 In other words we use interest rate values that are extreme and unlikely to be observed in a yield curve in practice. Bond traders would be amused if one morning they discovered that the bond redemption yield curve looked anything like Figure 7.10.

A graph of this extreme test data is set out in Figure 7.10.

The resulting smooth graph after the cubic spline parameter have been calculated and applied looks like that shown in Figure 7.11.

Two areas on the graph with relatively flat or consistent data values have been highlighted in Figure 7.11 as potential areas where calculation error may be observed. These areas of the graph are isolated and shown in Figures 7.12 and 7.13.

In the first area we observe some oscillation. However, this is not oscillation as a result of calculation errors. This is a smoothing effect that is required to meet the requirements of a cubic spline and to ensure a smooth curve. The data between points 63 and 71 is consistently downward sloping but the data then slopes upward again at point 72. The curve starts to 'adapt' at an

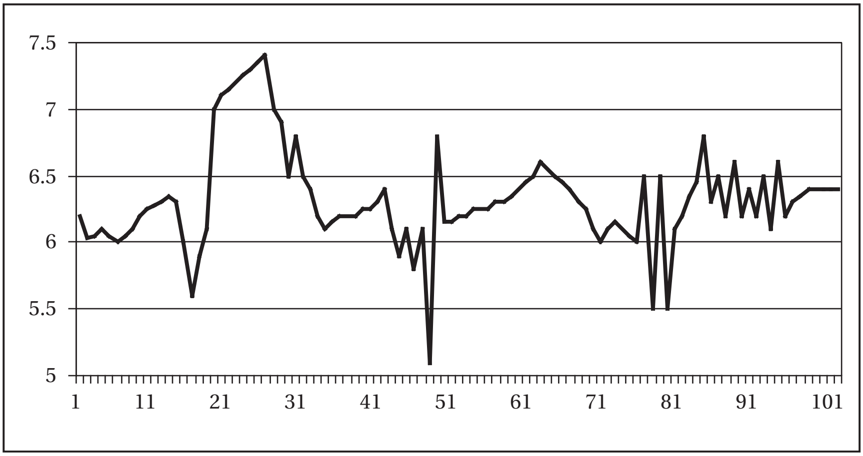


Figure 7.10: Graphic illustration of extreme test data.

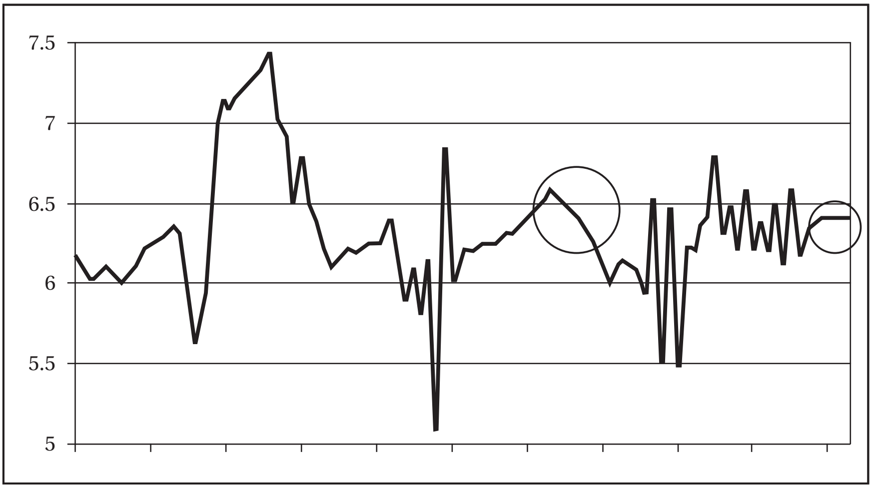


Figure 7.11: Graph from Figure 7.10 highlighting areas of error.

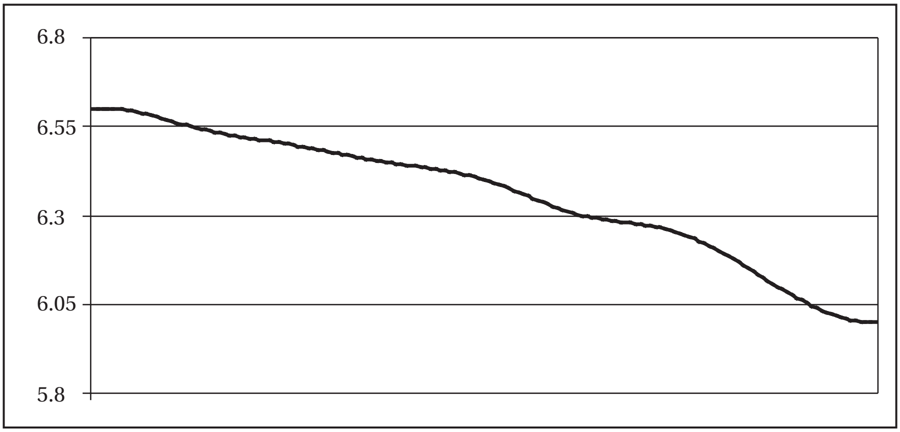


Figure 7.12: Exploded view of first error in Figure 7.11.

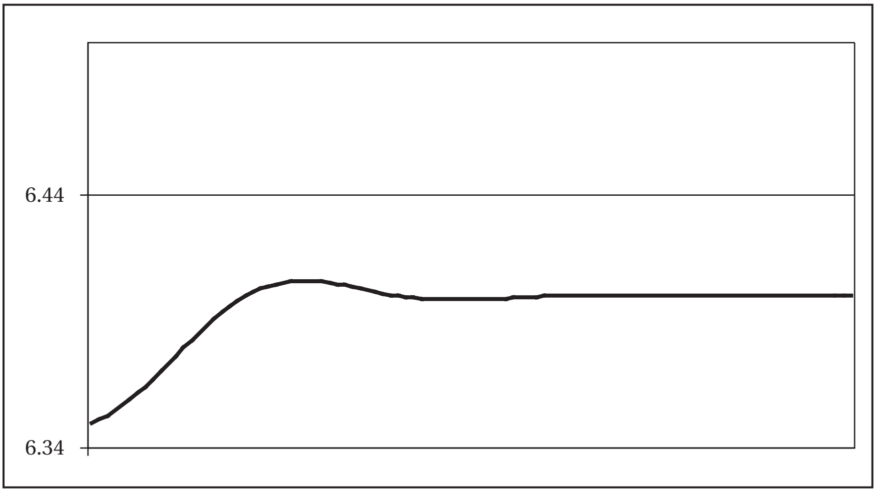


Figure 7.13: Exploded view of second error in Figure 7.11.

earlier stage in order to facilitate this change in direction. Therefore this behaviour is unavoidable, but under most applications for the spot curve does not present a material problem.

The second area of the curve provides another typical cubic spline example as the curve 'adapts' to its new parameters. Once again this is a natural spline phenomenon and not an error in the calculated values.

Empirical data does not prove beyond a doubt that a cubic spline method, applied using an appropriate solution technique and precise software, will always produce accurate results. Nonetheless we believe that it is reasonable to assume from the test data set out above that the cubic spline methodology, used in conjunction with appropriate calculation tools, provides accurate zero curve results in most fixed income market conditions.

7.6 A look at forward rates

Previous literature has highlighted the use of the cubic spline approach to model forward curves, and its limitations. Certainly a cubic spline discussion would be incomplete without a look at its application to forward rates. We will use our empirical data to highlight typical forward rate behaviour under the cubic spline technique. Our sample data does not reflect actual market conditions and is an extreme data set, to say the least. However, it does highlight a point with regards to forward rates that can often be observed sometimes under normal market conditions. To this end we isolate the last sub-set of the data, as shown above, and plot the forward rates for that data set.

From data that was interpolated using the linear method versus data interpolated using the cubic spline, a comparison of forwards shows how the forwards in a cubic spline environment can oscillate. As expected, the relatively minor oscillations observed first in the zero rates curve are compounded excessively in the forward rate calculation. The linear interpolation approach, shown for comparison purposes in Figure 7.14, eliminates much of the oscillation but of course is not a smooth curve, which is as undesirable. The user is confronted with the need to balance the conflicting requirements; a trade-off is called for and for most practical applications the cubic spline approach and its smoothing results is preferred. It remains important, however, that the user reviews cubic spline data by looking at both the zero and forward rates.

Using the actual United Kingdom 10-year zero curve for 2 January 2000, the forward rates have been calculated using cubic spline and linear interpolation and compared in Figures 7.15 and 7.16 respectively. There is no observed reason to favour the latter approach over the former.

7.6.1 *Improvements to the basic approach*

As a result of the drawback when fitting the forward curve, the basic technique has been improved to remove the oscillation effect at longer maturities. As we saw from the test results presented earlier, the oscillation of a spline is partly a function of the number of nodes used. The paradox with this factor is that in practice, at very long maturities the forward (and also the spot) curve would be expected to be reasonably flat. To remove the oscillation, as

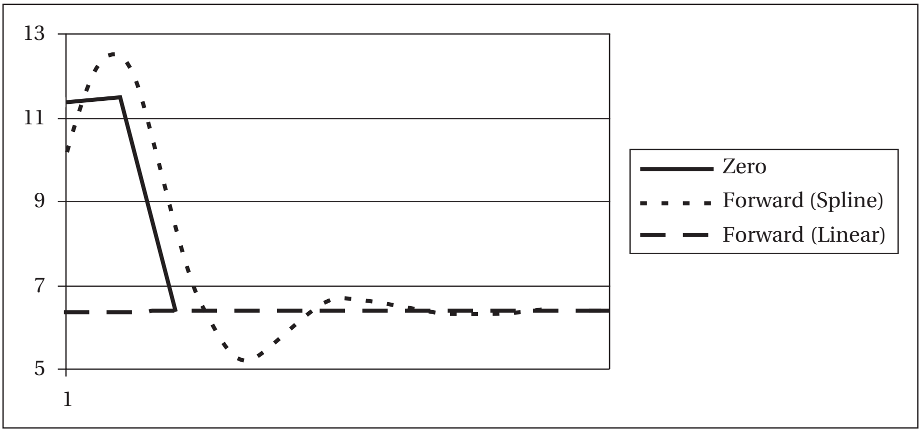


Figure 7.14: The linear approach, and unacceptable lack of 'smoothness'.

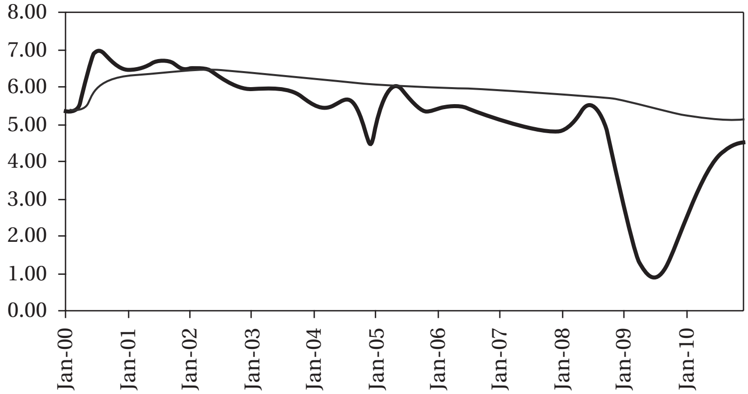


Figure 7.15: Graph prepared using cubic spline and linear interpolation.

described first by Fisher, Nychka and Zervos (1995), this involves the addition of a *roughness penalty* when minimising the sums of squares. Waggoner (1997) introduced a *variable roughness penalty*, which enabled the approach to retain the flexibility at the short end and reduce oscillation at the long end. Using the Waggoner approach enables users to retain the flexibility and ease of the cubic spline approach as well as a more realistic forward curve. Anderson and Sleath (1999) state that the advantage of the spline approach over parametric methods is that separate segments of the spline can be adjusted independently of each other. The significance of this is that a change in market levels at one of the term structure will not affect significantly at other parts of the curve. This is a drawback of the parametric methods. Ironically Anderson and Sleath modify the Waggoner model in a way that would appear to incorporate elements of the parametric approach, and their results appear to improve on the earlier works.

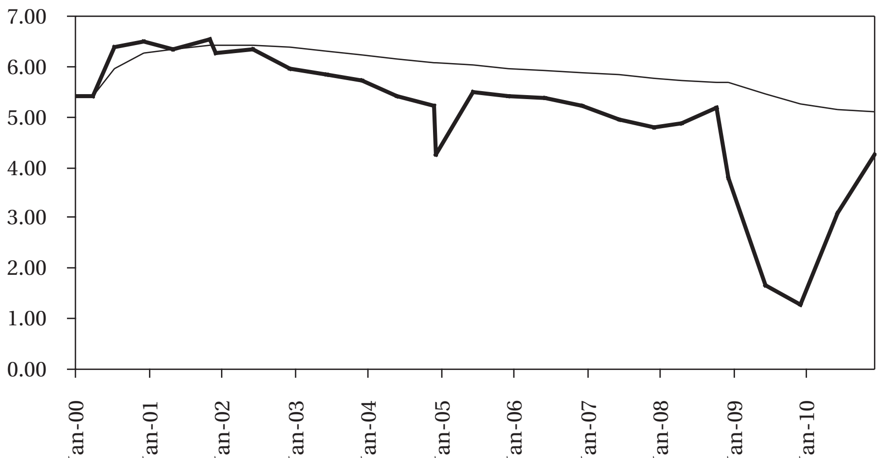


Figure 7.16: Graph prepared using cubic spline and linear interpolation.

7.7 Conclusion

The purpose of this chapter has been to present an accessible account of how the cubic spline methodology of term structure estimation could be implemented by users involved in any area of the debt capital markets. The technique is straightforward and quick, and is valid for a number of applications, most of which are 'normal' or conventional yield curves. For example users are recommended to use it when curves are positively sloping, or when the long end of the curve is not downward sloping. The existence of humps along the short or medium terms of the curve can cause excessive oscillation in the forward curve but the zero curve may still be used for valuation or relative value purposes.

Oscillation is a natural effect of the cubic spline methodology and its existence does not impair its effectiveness under many conditions. If observed rates produce very humped curves, the fitted zero-curve using cubic spline does not produce usable results. For policy making purposes, for example as used in central banks, and also for certain market valuation purposes, users require forward rates with minimal oscillation. In such cases, however, the Waggoner or Anderson–Sleath models will overcome this problem. The cubic spline approach can therefore be recommended under most market conditions.

Appendix

Appendix 7.1: Example matrix solution based on Gaussian elimination

We will solve for the following values (where the values of X have already been calculated).

x	X	y
0.90	0.40	1.30
1.30	0.60	1.50
1.90	0.20	1.85
2.10	0.90	2.10
3.00	0.80	1.95
3.80	0.50	0.40
4.30		0.25

Firstly we construct our matrix as follows.

$$
\left[\begin{array}{cccccc|c}
X_0 & 2(X_0+X_1) & X_1 & & & & -3\left(\frac{(d_1-d_0)}{X_0}-\frac{(d_2-d_1)}{X_1}\right) \\
& X_1 & 2(X_1+X_2) & X_2 & & & -3\left(\frac{(d_2-d_1)}{X_1}-\frac{(d_3-d_2)}{X_2}\right) \\
& & \dots & \dots & \dots & & \dots \\
& & & X_{N-2} & 2(X_{N-2}+X_{N-1}) & X_{N-1} & -3\left(\frac{(d_{N-1}-d_{N-2})}{X_{N-2}}-\frac{(d_N-d_{N-1})}{X_{N-1}}\right)
\end{array}\right]
$$

Where b_1 is set to zero this provides the values.

b1	b2	b3	b4	b5	b6	b7	
0.0	2.0	0.6					0.3
	0.6	1.6	0.2				2.0
		0.2	2.2	0.9			−4.3
			0.9	3.4	0.8		−5.3
				0.8	2.6	0.5	4.9

In turn we can substitute row 1 into row 2 to obtain.

b1	b2	b3	b4	b5	b6	b7	
0.0	2.0	0.6					0.3
	0.0	4.7	0.7				6.4
		0.2	2.2	0.9			−4.3
			0.9	3.4	0.8		−5.3
				0.8	2.6	0.5	4.9

Similar substitutions, and the fact that b_7 is constrained as zero, yield the matrix below.

b1	b2	b3	b4	b5	b6	b7	
0.0	2.0	0.6					0.3
	0.0	4.7	0.7				6.4
		0.0	51.4	21.3			−107.0
			0.0	172.9	45.7		−196.4
				0.0	516.2	0.0	1,258.0

This means that we can solve for b_6. Once we have a solution for b_6 we can solve for b_5 and so on. As a final result we get the following values for parameter b.

b1	b2	b3	b4	b5	b6	b7
0.0	−0.338	1.544	−1.344	−1.780	2.437	0.0

Parameters a and c can be determined directly from the values of b above.

Selected bibliography and references

Adams, K.J., van Deventer, D.R., 'Fitting yield curves and forward rate curves with maximum smoothness', *Journal of Fixed Income* 4, 1994, 52–62.

Anderson, N., Sleath, J., 'New estimates of the UK real and nominal yield curves', *Bank of England Quarterly Bulletin*, November 1999, pp. 384–392.

Beim, D., 'Term structure and the non-cash value in bonds', *First Boston Working Paper series*, 1992.

Bjork, T., Christensen, B., 'Interest rate dynamics and consistent forward rate curves', *University of Aarhus Working Paper*, 1997, pp. 1–38.

Burden, R., Faires, D., *Numerical Analysis*, Brooks/Cole, 1997.

De la Grandville, O., *Bond Pricing and Portfolio Analysis*, MIT Press, 2001, pp. 248–252.

Fisher, M., Nychka, D., Zervos, D., 'Fitting the term structure of interest rates with smoothing splines', Working Paper No. 95-1, *Finance and Economics Discussion Series*, Federal Reserve Board, 1995.

James, J., Webber, N., *Interest Rate Modelling*, Wiley, 2000, p. 444.

McCulloch, J., 'The tax-adjusted yield curve', *Journal of Finance* 30, 1975, pp. 811–830.

Nelson, C., Siegel, A., 'Parsimonious modeling of the yield curve', *Journal of Business* 60, no. 4, 1987, pp. 473–489.

Svensson, L., 'Estimating forward interest rates with the extended Nelson and Siegel method', *Sveriges Riksbank Quarterly Review* 3, 1995.

Vasicek, O., Fong, H.G., 'Term structure modelling using exponential splines', *Journal of Finance* 37 (2), May 1982, pp. 339–356.

Waggoner, D., 'Spline methods for extracting interest rate curves from coupon bond prices', *Working Paper No. 97-10*, Federal Reserve Bank of Atlanta, 1997.

Wiseman, J., 'The exponential yield curve model', *JPMorgan European Fixed Income Research*, 1994.

8 Advanced Analytics for Index-linked Bonds

Bonds that have part or all of their cash flows linked to an inflation index form an important segment of several government bond markets. In the United Kingdom the first index-linked bonds were issued in 1981 and at the end of 1999 they accounted for approximately 15% of outstanding nominal value in the gilt market. Index-linked bonds were only recently introduced in the US Treasury market but are more established in Australia, Canada, the Netherlands, New Zealand and Sweden. There is no uniformity in market structure and as such there are significant differences between the index-linked markets in these countries. There is also a wide variation in the depth and liquidity of these markets.

Index-linked bonds or inflation-indexed bonds present additional issues in their analysis, due to the nature of their cash flows. Measuring the return on index-linked bonds is less straightforward than with conventional bonds, and in certain cases there are peculiar market structures that must be taken into account as well. For example, in the US market for index-linked Treasuries (known as 'TIPS' from Treasury inflation-indexed securities) there is no significant lag between the inflation link and the cash flow payment date. In the UK there is an eight-month lag between the inflation adjustment of the cash flow and the cash flow payment date itself, while in New Zealand there is a three-month lag. The existence of a lag means that inflation protection is not available in the lag period, and that the return in this period is exposed to inflation risk; it also must be taken into account when analysing the bond.

From market observation we know that index-linked bonds can experience considerable volatility in prices, similar to conventional bonds, and therefore there is an element of volatility in the real yield return of these bonds. Traditional economic theory states that the level of real interest rates is constant, however in practice they do vary over time. In addition there are liquidity and supply and demand factors that affect the market prices of index-linked bonds. In this chapter we present analytical techniques that can be applied to index-linked bonds, the duration and volatility of index-linked bonds and the concept of the real interest rate term structure.

8.1 Index-linked bonds and real yields

The real return generated by an index-linked bond, or its real yield, is usually defined as yield on risk-free index-linked bonds, or in other words the long-term interest rate on risk-free funds minus the effect of inflation. There may also be other factors involved, such as the impact of taxation. Therefore the return on an index-linked bond should in theory move in line with the real cost of capital. This will be influenced by the long-term growth in the level of real gross domestic product in the economy. This is because in an economy experiencing rapid growth, real interest rates are pushed upwards as the demand for capital increases, and investors therefore expect higher real yields. Returns are also influenced by the demand for the bonds themselves.

The effect of general economic conditions and the change in these over time results in real yields on index-linked bonds fluctuating over time, in the same way nominal yields fluctuate for conventional bonds. This means that the price behaviour of indexed bonds can also be fairly volatile.

The yields on indexed bonds can be used to imply market expectations about the level of inflation. For analysts and policy makers to use indexed bond yields in this way, it is important that a liquid secondary market exists in the bonds themselves. For example the market in Australian index-linked bonds is relatively illiquid, so attempting to extract an information content from their yields may not be valid. Generally though the real yields on indexed bonds reflect investors' demand for an inflation premium, or rather a premium for the uncertainty regarding future inflation levels. This is because holders of indexed bonds are not exposed to inflation-eroded returns; therefore if future inflation was expected to be zero, or known with certainty (whatever its level) there would be no requirement for an inflation premium, because there would be no uncertainty. In the same way, the (nominal) yields on conventional bonds reflect market expectations on inflation levels. Therefore higher volatility of the expected inflation rate will lead to a higher inflation risk premium on conventional bonds, and a lower real yield on indexed bonds relative to nominal yields. It is the uncertainty regarding future inflation levels that creates a demand for an inflation risk yield premium, rather than past experience of inflation levels. However investor sentiment may well demand a higher inflation premium in a country with a poor record in combating inflation.

Traditionally information on inflation expectations has been obtained by survey methods or theoretical methods. These have not proved reliable however, and were followed only because of the absence of an inflation-indexed futures market.[1] Certain methods for assessing market inflation expectations are not analytically valid; for example the suggestion that the spread between short-term and long-term bond yields cannot be taken to be a measure of inflation expectation, because there are other factors that drive this yield spread, and not just inflation risk premium. Equally, the spread between the very short-term (overnight or one week) interest rate and the two-year bond yield cannot be viewed as purely driven by inflation expectations. Using such approaches to glean information on inflation expectations is logically unsound. One approach that is valid, as far as it goes, would be to analyse the spread between historical real and nominal yields, although this is not a forward-looking method. It would however indicate the market's inflation premium over a period of time. The best approach though is to use the indexed bond market; given a liquid market in conventional and index-linked bonds it is possible to derive estimates of inflation expectations from the yields of both sets of bonds. This is reviewed later in the chapter.

8.2 Duration and index-linked bonds

In earlier chapters we reviewed the basic features of index-linked bonds and their main uses. We also discussed the techniques used to measure the yield on these bonds. The

[1] The New York Coffee, Sugar and Cocoa Exchange traded a futures contract on the US consumer prices index (CPI) in the 1980s.

largest investors in indexed bonds are long-dated institutions such as pension fund managers, who use them to match long-dated liabilities that are also index-linked, for example a pension contract that has payments linked to the inflation index. It is common though for investors to hold a mixture of indexed and conventional bonds in their overall portfolio.

The duration of a bond is used as a measure of its sensitivity to changes in interest rate. The traditional measure, if applied to indexed bonds, will result in high values due to the low coupon on these bonds and the low real yield. In fact the longest-duration bonds in most markets are long-dated indexed bonds. The measure, if used in this way however, is not directly comparable to the duration measure for a conventional bond. Remember that the duration of a conventional bond measures its sensitivity to changes in (nominal) yields, or put another way to changes in the combined effect of inflation expectations and real yields. The duration measure of an indexed bond on the other hand, would be a measure of its sensitivity to changes in real yields only, that is to changes in real interest rate expectations. Therefore it is not valid to compare traditional duration measures between conventional bonds and indexed bonds, because one would not be comparing like-for-like. This has important implications for the portfolio level. If a portfolio is composed of both conventional and indexed bonds, how does one measure its combined duration? The traditional approach of combining the duration values of individual bonds would have no meaning in this context, because the duration measure for each type of bond is measuring something different. For example, consider a situation where there are two portfolios with the same duration measure. If one portfolio was composed of a greater amount by weighting of index-linked bonds, it would have a different response to changes in market yields, especially so if investors' economic expectations shifted significantly, compared to the portfolio with a lower weighting in indexed bonds. Therefore a duration-based approach to market risk would no longer be adequate as a means of controlling portfolio market risk.

Therefore the key focus of fund managers that run combined portfolios of conventional and indexed bonds is to manage the duration of the conventional and indexed bonds on a separate basis, and to be aware of the relative weighting of the portfolio in terms of the two bond types. A common approach is to report two separate duration values for the portfolio, which would measure two separate types of risk exposure. One measure would be the *portfolio real yield duration*, which is the value of the combined durations of both the conventional and indexed bonds. This measure is an indication of how the portfolio value will be affected by a change in market real yields, which would impact both indexed and conventional bond yields. The other measure would be the *portfolio inflation duration*, which is a duration measure for the conventional bonds only. This duration measure indicates the sensitivity of the portfolio to a change in market inflation expectations, which have an impact on nominal yields but not real yields. Portfolio managers also follow a similar approach with regard to interest-rate volatility scenarios. Therefore if carrying out a parallel yield curve shift simulation, which in terms of a combined portfolio would actually correspond to a real-yield simulation, the portfolio manager would also need to undertake a simulation that mirrored the effect of a change in inflation expectations, which would have an impact on nominal yields only.

The traditional duration approach can be used with care in other areas. For instance, the Bank of England monetary policy committee is tasked with keeping inflation at a level of 2.5%. If therefore the ten-year benchmark gilt is trading at a yield of 6.00% while the ten-year index-linked gilt is trading at a real yield of 3.00%, this implies that the market expectation of average inflation rates during the next ten years is 3.00%. This would suggest that the benchmark gilt is undervalued relative to the indexed gilt. To effect a trade that matched the market maker's view, one would short the ten-year index-linked gilt and buy the conventional gilt. If the view turned out to be correct and market inflation expectations declined, the trade would generate a profit. If on the other hand real interest-rate expectations changed, thus altering real yields, there would be no effect. The other use of the traditional duration approach is with regard to hedging. Indexed bonds are sometimes difficult to hedge because of the lack of suitable hedging instruments. The most common hedging instrument is another indexed bond, and the market maker would use a duration weighting approach to calculate the nominal value of the hedging bond.

In the traditional approach the duration value is calculated using nominal cash flows, discounted at the nominal yield. A more common approach is to assume a constant average rate of inflation, and adjust cash flows using this inflation rate. The real yield is then used to discount the assumed future cash flows. There are a number of other techniques that can be used to calculate a duration value, all requiring the forecasting of the level of future cash flows and discounting using the nominal yield. These include:

- as above, assuming a constant average inflation rate, which is then used to calculate the value of the bond's coupon and redemption payments. The duration of the cash flow is then calculated by observing the effect of a parallel shift in the zero-coupon yield curve. By assuming a constant inflation rate and constant increase in the cash flow stream, a further assumption is made that the parallel shift in the yield curve is as a result of changes in real yields, not because of changes in inflation expectations. Therefore this duration measure becomes in effect a real yield duration;
- a repeat of the above procedure, with the additional step, after the shift in the yield curve, of recalculating the bond cash flows based on a new inflation forecast. This produces a duration measure that is a function of the level of nominal yields. This measure is in effect an inflation duration, or the sensitivity to changes in market inflation expectations, which is a different measure to the real yield duration;
- an assumption that the inflation scenario will change by an amount based on the historical relationship between nominal yields and the market expectation of inflation. This is in effect a calculation of nominal yield duration, and would be a measure of sensitivity to changes in nominal yields.

Possibly the most important duration measure is the real yield duration, which is more significant in markets where there is a lag between the indexation and cash flow dates, due to the inflation risk exposure that is in place during the lag period. This is the case in both the UK and Australia, although as we noted the lag is not significant in the US market. It is worth noting that index-linked bonds do not have stable nominal duration values, that is, they do not exhibit a perfectly predictable response to changes in nominal yields. If they did, there would be no advantage in holding them, as their behaviour could be replicated by conventional bonds. For this reason, index-linked bonds cannot be hedged perfectly with conventional bonds, although this does happen in practice on occasions when no other hedging instrument is available.

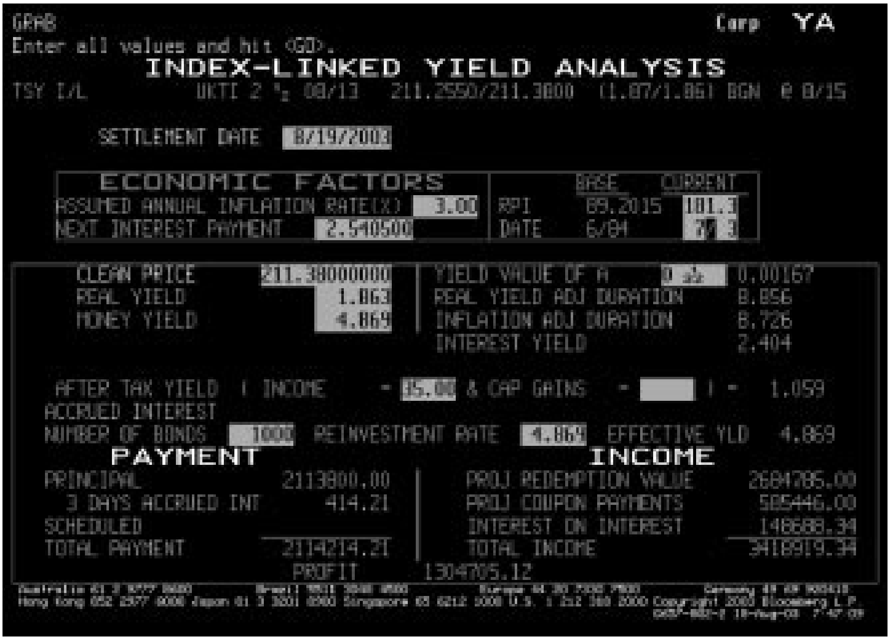


Figure 8.1: Example of index-linked yield analysis, UK 2½% Treasury 2009 (assumed annual inflation rate 3.00%, base inflation index 89.2015, current index 181.3), showing real yield and money yield, 18 August 2003. ©Bloomberg L.P. Reproduced and used with permission.

One final point regarding duration is that it is possible to calculate a *tax-adjusted duration* for an index-linked bond in markets where there is a different tax treatment to indexed bonds compared to conventional bonds. In the US market the returns on indexed and conventional bonds are taxed in essentially the same manner, so that in similar fashion to Treasury strips, the inflation adjustment to the indexed bond's principal is taxable as it occurs, and not only on the maturity date. Therefore in the US indexed bonds do not offer protection against any impact of after-tax effects of high inflation. That is, Tips real yields reflect a premium for only pre-tax inflation risk. In the UK market however, index-linked gilts receive preferential tax treatment, so their yields also reflect a premium for after-tax inflation risk. In practice this means that the majority of indexed gilt investors are those with high marginal tax rates.[2] This factor also introduces another element in analysis; if the demand for indexed or conventional bonds were to be a function of expected after-tax returns, this would imply that pre-tax real yields should rise as expected inflation rates rise, in order to maintain a constant after-tax real yield. This has not been observed explicitly in practice, but is a further factor of uncertainty about the behaviour of real yields on index-linked bonds.[3]

2 For example, see Brown and Schaefer (1996).

3 For further detail on this phenomenon, see Roll (1996).

8.3 Estimating the real term structure of interest rates

In Chapter 11 of the author's book *The Bond and Money Markets* we show some approaches used to measure inflation expectations, with reference to UK index-linked gilts. To recap, these measures include:

- the 'simple' approach, where the average expected inflation rate is calculated using the Fisher identity, so that the inflation estimate is regarded as the straight difference between the real yield on an index-linked bond, at an assumed average rate of inflation, and the yield on a conventional bond of similar maturity;
- the 'break-even' inflation expectation, where average inflation expectations are estimated by comparing the return on a conventional bond against that on an indexed bond of similar maturity, but including an application of the compound form of the Fisher identity. This has the effect of decomposing the nominal rate of return on the bond into components of real yield and inflation;
- a variation of the break-even approach, but matching stocks by duration rather than by maturity.

The drawbacks of each of these approaches are apparent. A rather more valid and sound approach is to construct a term structure of the real interest rates, which would indicate, in exactly the same way that the conventional forward rate curve does for nominal rates, the market's expectations on future inflation rates. In countries where there are liquid markets in both conventional and inflation-indexed bonds, we can observe a nominal yield curve and a real yield curve. It then becomes possible to estimate both a conventional term structure and a real term structure; using these allows us to create pairs of hypothetical conventional and indexed bonds that have identical maturity dates, for any point on the term structure.[4] We could then apply the break-even approach to any pair of bonds to obtain a continuous curve for both the average and the forward inflation expectations. To maximise use of the available information we can use all the conventional and indexed bonds that have reasonable liquidity in the secondary market.

In this section we review one method that can be used to estimate and fit a real term structure.

8.3.1 *The term structure of implied forward inflation rates*

In previous chapters we reviewed the different approaches to yield curve modelling used to derive a nominal term structure of interest rates. We saw that the choice of yield curve model can have a significant effect on the resulting term structure; in the same way, the choice of model will impact the resulting real rate term structure as well. One approach has been described by McCulloch (1975), while in the UK market the Bank of England uses a modified version of the approach posited by Waggoner (1997) which we discussed in the previous chapter. McCulloch's approach involves estimating a discount function by imposing a constraint on the price of bonds in the sample to equal the sum of the

---

4 We are restricted however to the longest-dated maturity of either of the two types of bonds.

discounted values of the bonds' cash flows. The Waggoner approach uses a cubic spline-based method, like McCulloch, with a roughness penalty that imposes a trade-off between the smoothness of the curve and its level of forward rate oscillation. The difference between the two approaches is that with McCulloch it is the discount function that is specified by the spline function, whereas in the Waggoner model it is the zero-coupon curve. Both approaches are valid, in fact due to the relationship between the discount function, zero-coupon rate and forward rate, both methods will derive similar curves under most conditions.

Using the prices of index-linked bonds it is possible to estimate a term structure of real interest rates. The estimation of such a curve provides a real interest counterpart to the nominal term structure that was discussed in the previous chapters. More important it enables us to derive a real forward rate curve. This enables the real yield curve to be used as a source of information on the market's view of expected future inflation. In the UK market there are two factors that present problems for the estimation of the real term structure; the first is the eight-month lag between the indexation uplift and the cash flow date, and the second is the fact that there are fewer index-linked bonds in issue, compared to the number of conventional bonds. The indexation lag means that in the absence of a measure of expected inflation, real bond yields are dependent to some extent on the assumed rate of future inflation. The second factor presents practical problems in curve estimation; in December 1999 there were only 11 index-linked gilts in existence, and this is not sufficient for most models. Neither of these factors presents an insurmountable problem however, and it is still possible to estimate a real term structure.

8.3.2 *Estimating the real term structure*[5]

There are a number of techniques that can be applied in estimating the real term structure. One method was described by Schaefer (1981). The method we describe here is a modified version of the cubic spline technique described by Schaefer. This is a relatively straightforward approach. The adjustment involves simplifying the model, ignoring tax effects and fitting the yield-to-maturity structure. A reduced number of nodes defining the cubic spline is specified compared with the conventional term structure, because of the fewer number of index-linked bonds available, and usually only three node points are used. Our approach therefore estimates three parameters, defining a spline consisting of two cubic functions, using 11 data points. The approach is defined below.

In the first instance, we require the real redemption yield for each of the indexed bonds. This is the yield that is calculated by assuming a constant average rate of inflation, applying this to the cash flows for each bond, and computing the redemption yield in the normal manner. The yield is therefore the market-observed yield, using the price quoted for each bond. These yields are used to define an initial estimate of the real yield curve, as they form the initial values of the parameters that represent the real yield at each node point. The second step is to use a non-linear technique to estimate the values of the parameters that will minimise the sum of the squared residuals between

5 This section follows the approach (with permission) from Deacon and Derry (1994), a highly accessible account. This is their Bank of England working paper, 'Deriving Estimates of Inflation Expectations from the Prices of UK Government Bonds'.

the observed and fitted real yields. The fitted yield curve is viewed as the real par yield curve; from this curve we calculate the term structure of real interest rates and the implied forward rate curve, using the technique described in Chapter 6. In estimating the real term structure in this way, we need to be aware of any tax effects. In the UK market, there is a potentially favourable tax effect, which may not apply in say, the US Tips market. Generally for UK indexed gilts, high marginal taxpayers are the biggest holders of index-linked bonds because of the ratio of capital gain to income, and their preference is to hold shorter-dated indexed bonds. On the other hand pension funds, which are exempt from income tax, prefer to hold longer-dated indexed gilts. The approach we have summarised here ignores any tax effects, but to be completely accurate any tax impact must be accounted for in the real term structure.

8.3.3 *Fitting the discount function*

The term structure method described by McCulloch (1971) involved fitting a discount function, rather than a spot curve, using the market prices of a sample of bonds. This approach can be used with only minor modifications to produce a real term structure.

Given the bond price equation (8.1):

$$P_i = C_i \int_0^{T_i} \mathrm{d}f(\mu)\mathrm{d}\mu + M_i \mathrm{d}f(T_i) \tag{8.1}$$

where P_i, C_i, T_i, M_i are the price, coupon, maturity and principal payment of the ith bond, we set the set of discrete discount factors as the discount function df, defined as a linear combination of a set of k linearly independent underlying basis functions, given by (8.2):

$$\mathrm{d}f(T) = 1 + \sum_{j=1}^{k} a_j f_j(T) \tag{8.2}$$

where $f_j(T)$ is the jth basis function and a_j is the corresponding coefficient, with $j = 1, 2, \ldots, k$. It can be shown (see Deacon and Derry (1994)) that for index-linked bonds equation (8.2) can be adapted by a scaling factor Δ_i that is known for each bond, once an assumption has been made about the future average inflation rate, to fit a discount function for indexed bonds. We estimate the coefficients a_j from:

$$y_i = \sum_{j=1}^{k} a_j x_{ij}$$

where

$$y_i = P_i - \Delta_i C_i T_i - \Delta_i M_i$$

$$x_{ij} = \Delta_i C_i \int_0^{T_i} f_j \mu \mathrm{d}\mu + \Delta_i M_i f_j(T_i)$$

$$u = (1 + \pi^e)^{-1/2}$$

$$\Delta_i = \begin{cases} [u^{\mathrm{td}j}]_i \cdot \dfrac{RPID_j}{RPIB_i} & \text{if } RPID_j \text{ is known} \\ [u^{\mathrm{td}l-L/6}]_i \cdot \dfrac{RPIL}{RPIB_i} & \text{otherwise} \end{cases}$$

where P_i, C_i, T_i, M_i are as before, but this time representing the index-linked bond. The scaling factor Δ_i is that for the ith bond, and depends on the ratio of the retail price index (RPI) at the time compared to the RPI level in place at the time the bond was issued, known as the *base* RPI.[6] If in fact the RPI that is used to index any particular cash flow is not known, it must be estimated using the latest available RPI figure, in conjunction with an assumption about the path of future inflation, using π^e.

8.3.4 *Deriving the term structure of inflation expectations*

Using any of the methods described in Chapter 6 or the discount function approach summarised above, we can construct curves for both the nominal and the real implied forward rates. These two curves can then be used to infer market expectations of future inflation rates. The term structure of forward inflation rates is obtained from both these curves by applying the Fisher identity:

$$1 + \frac{f}{2} = (1+i)^{1/2}\left(1 + \frac{r}{2}\right) \tag{8.3}$$

where f is the implied nominal forward rate, r is the implied real forward rate and i is the implied forward inflation rate. As with the term structure of real spot rates, the real implied forward rate curve is dependent on an assumed rate of inflation. To make this assumption consistent with the inflation term structure that is calculated, we can use an iterative procedure for the assumed inflation rate. Essentially this means that the real yield curve is re-estimated until the assumed inflation term structure and the estimated inflation term structure are consistent. Real yields are usually calculated using either a 3% or a 5% flat inflation rate. This enables us to estimate the real yield curve, from which the real forward rate curve is derived. Using (8.3) we can then obtain an initial estimate of the inflation term structure. This forward inflation curve is then converted into an average inflation curve, using (8.4):

$$i_i = \prod_j^k (1 + if_i)^{-1/k} - 1 \tag{8.4}$$

where

if_i is the forward inflation rate at maturity i

i_i is the average inflation rate at maturity i.

From this average inflation curve, we can select specific inflation rates for each index-linked bond in our sample. The real yields on each indexed bond are then re-calculated

---

6 Due to the lag in the UK gilt market, for index-linked gilts the base RPI is actually the level recorded for the month eight months before the issue date.

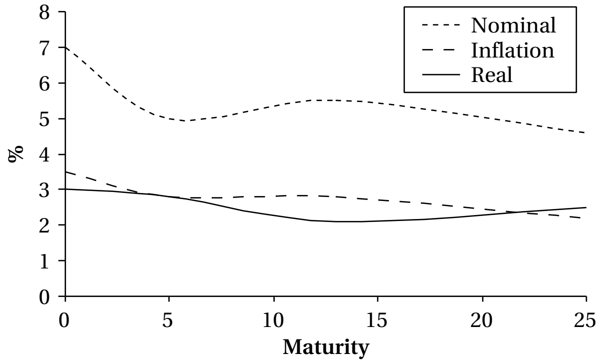


Figure 8.2: UK market nominal and real term structure of interest rates, July 1999. Yield source: BoE.

using these new inflation assumptions. From these yields the real forward curve is calculated, enabling us to produce a new estimate of the inflation term structure. This process is repeated until there is consistency between the inflation term structure used to estimate the real yields and that produced by (8.3).

Using the modified Waggoner method described in Chapter 6, the nominal spot yield curve for the gilt market in July 1999 is shown in Figure 8.2. The real term structure is also shown, which enables us to draw the implied forward inflation expectation curve, which is simply the difference between the first two curves.

8.3.5 *Application*

Real yield curves are of some use to investors, for a number of reasons. These include applications that arise in insurance investment management and corporate finance, such as the following:

- they can be used to value inflation-linked liabilities, such as index-linked annuity contracts;
- they can be used to value inflation-linked revenue streams, such as taxes that are raised in line with inflation, or for returns generated in corporate finance projects; this makes it possible to assess the real returns of project finance or government revenue;
- they can be used to estimate the present value of a company's future staff costs, which are broadly linked to inflation.

Traditionally, valuation methods for such purposes would use nominal discount rates and an inflation forecast, which would be constant. Although the real term structure also includes an assumption element, using estimated market real yields is equivalent to using a nominal rate together with an implied market inflation forecast, which need not be constant. This is a more valid approach; a project financier in the UK in July 1999 can obtain more meaningful estimates on the effects of inflation using the rates implied in Figure 8.2, rather than an arbitrary, constant inflation rate. The inflation term structure can be used in other ways as well; for example, an investor in mortgage-backed bonds, who uses a prepayment model to assess the prepayment risk associated with the bonds, will make certain assumptions about the level of prepayment of the mortgage pool backing the bond. This prepayment rate is a function of a number of factors, including the level of interest rates, house prices and the general health of the economy. Rather than use an arbitrary assumed prepayment rate, the rate can be derived from market inflation forecasts.

In essence, the real yield curve can and should be used for all the purposes for which the nominal yield curve is used. Provided that there are enough liquid index-linked bonds in the market, the real term structure can be estimated using standard models, and the result is more valid as a measure of market inflation expectations than any of the other methods that have been used in the past.

Selected bibliography and references

Arak, M., Kreicher, L., 'The real rate of interest: inferences from the new UK indexed gilts', *International Economic Review* 26(2), June 1985, pp. 399–408.

Bootle, R., *Index-Linked Gilts*, 2nd edition, Woodhead-Faulkner, 1991.

Brown, R., Schaefer, S., 'Ten years of the real term structure: 1984–1994', *Journal of Fixed Income*, March 1996, pp. 129–141.

Brynjolfsson, J., Fabozzi, F. (eds.), *Handbook of Inflation-Indexed Bonds*, FSF Associates, 1999.

Deacon, M., Derry, A., 'Deriving estimates of inflation expectations from the prices of UK government bonds', *Bank of England Working Paper* No. 23, 1994.

Deacon, M., Derry, A., *Inflation-Indexed Securities*, Prentice Hall, 1998.

McCulloch, J., 'Measuring the term structure of interest rates', *Journal of Business* XLIV, January 1971, pp. 19–31.

McCulloch, J., 'The tax-adjusted yield curve', *Journal of Finance* 30(3), June 1975, pp. 811–830.

Roll, R., 'US Treasury inflation-indexed bonds: the design of a new security', *Journal of Fixed Income*, December 1996, pp. 321–335.

Schaefer, S., 'Measuring a tax-specific term structure of interest rates in the market for British government securities', *The Economic Journal* 91, June 1981, pp. 415–438.

Waggoner, D., 'Spline methods for extracting interest rate curves from coupon bond prices', *Working Paper*, No. 97–10, Federal Reserve Bank of Atlanta, 1997.

This page intentionally left blank

9 Analysing the Long Bond Yield

A common observation in government bond markets is that the longest-dated bond trades expensive to the yield curve. It also exhibits other singular features that have been the subject of recent research for example by Phoa (1998), which we review in this chapter. The main feature of long bond yields is that they reflect a convexity effect. Analysts have attempted to explain the convexity effects of long bond yields in a number of ways. These are discussed first. We then consider the volatility and convexity bias that is observed in long bond yields.

9.1 Theories of long-dated bond yields

In both the United States and United Kingdom government markets extremely long-dated bonds have been actively traded. These include 'century' bonds in the US and undated or irredeemable bonds in the UK, including gilt issues such as the 'consols' or War Loan. At what yields should very long-dated bonds trade? Under the conventional hypothesis reviewed in Chapter 2, investors might believe that if the yield on the 30-year government bond is 6.00%, the yield on a hypothetical 100-year government bond will be higher, say 6.25%, the higher rate signifying the term premium payable on the longer bond. In fact, this is extremely unlikely, and it has been shown that, for such a term structure to be observed, we would require forward interest rates to be very high. Expected rates will not therefore be as high as the forward rates. We explore the issues in this section.

Theories of very long-dated interest rates have been proposed that on observation, would appear to hold; these include:

- that very long-dated yields are not an unbiased average of expected future interest rates, but rather can be estimated using a weighting of various interest-rate scenarios; at sufficiently long maturities the highest interest-rate scenarios do not impact the long-dated yield (Dybvig and Marshall 1996);
- extremely long-dated zero-coupon and forward rates can never decline, even when expected long-term future interest rates fall; therefore this limits the extent to which very long-dated bond yields are affected by a change in the current interest-rate environment (Dybvig, Ingersoll and Ross 1996).

The very long-dated zero-coupon yield is taken to be the infinite maturity zero-coupon yield, that is the limiting yield of a risk-free zero-coupon bond whose maturity approaches infinity. Although it might appear so, the infinite maturity yield is not identical to the yield on an irredeemable bond, which pays coupons during its life and so has a shorter-dated yield. It is also not identical to the long-term interest rate, which is defined as the expected long-term rate of return on bonds, or the expected rate of return on a bond with infinite duration. The long rate is a measure of the expected future rate of *return*, rather than a present bond *yield*. The two interest-rate hypotheses above are general and apply to both conventional and index-linked bonds. They use the principal of no-arbitrage pricing, in terms of a trading strategy, in their derivation, which we do not present here. They do however have a practical significance in terms of the valuation of long-dated bonds.

9.1.1 *Long-dated yields*

In an environment of interest-rate uncertainty, from previous chapters we know the price today of a zero-coupon bond of maturity T to be a function of the expectation of future short rates, which at time t are not known; this is given in (9.1):

$$P(t,\ T) = exp\left(-\int_t^T r(s)\mathrm{d}s\right). \tag{9.1}$$

Expression (9.1) states that the price of a zero-coupon bond is equal to the discount factor from time t to its maturity date, or the average of the discount factors under all interest-rate scenarios, weighted by their probabilities. It can be shown that the T-maturity forward rate at time t is given by

$$f(t,\ T) = \frac{E_t\left[r_T\ exp\left(-\int_t^T r(s)ds\right)\right]}{E_t\left[exp\left(-\int_t^T r(s)ds\right)\right]} \tag{9.2}$$

which expresses the T-term forward rate in terms of the dynamics of the T-maturity short-rate r_T under all possible interest-rate scenarios, that is, along all possible random interest-rate paths. The weightings are in terms of the probabilities of each interest-rate path occurring and the discount factors from the period t to T that occur in each scenario. So for instance consider an environment where there are only two possible interest-rate scenarios, each with a probability of $p(1)$ and $p(2)$. Phoa (1998) states that the T-maturity forward rate is given by the weighted average, shown in (9.3) where Df is the discount factor:

$$f(t,\ T) = \frac{p(1)Df_1(t,\ T)\cdot r_{T,(1)} + p(2)Df_2(t,\ T)\cdot r_{T,(2)}}{p(1)Df_1(t,\ T) + p(2)Df_2(t,\ T)}. \tag{9.3}$$

The effect of weighting using discount factors is to make the lower level interest-rate scenario more significant, because the discount factors are higher under these scenarios. This means that a lower interest-rate scenario has more influence on the forward rate than a higher-rate scenario, and this influence steadily increases as the forward rate term grows in maturity, since the difference between the discount factors increases. This is an important result.

9.1.2 *Long-dated forward rates*

Following Phoa (1998), we can illustrate this with a hypothetical example. Consider a binomial interest-rate environment under which there are the following interest-rate scenarios:

- the short-rate is at a constant level of 8.00%, with a probability of 70%;
- the short-rate is at a constant level of 4.00%, with a probability of 30%.

The expected future short-rate at any point in the future will be 8%, given the probabilities, however the forward rate will be lower than 8%, because it is calculated by weighting each interest-rate scenario by the relevant discount factors. This is illustrated in Figure 9.1.

The weight attached to the lower 4% interest-rate scenario increases with increasing term-to-maturity, while the weight on the higher rate will diminish. Therefore the 30-year

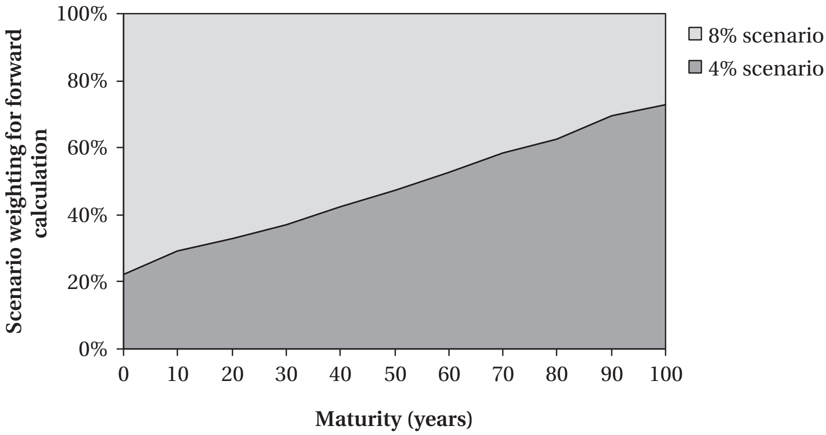


Figure 9.1: Forward rate calculation weighted by discount factors.

forward rate will be below 8% while the 100-year forward rate will be around half of the short-rate for the most probable scenario. Put another way, over a long period only the lowest interest-rate scenario is relevant, which is the theorem posited by Dybvig and Marshall. This tendency for the forward rate curve to fall with very long maturities is as a result of a *convexity bias* in the behaviour of the yield curve, which we consider later. This effect influences zero-coupon yields, which also exhibit a tendency to gravitate towards the lowest interest-rate scenario. Consider now another hypothetical example, where the current short-rate is 6%, and that there are now three (and only three) possible interest-rate scenarios, which are:

- that the short-rate increases from 6% to a long-term rate of 10%;
- that the short-rate increases from 6% to a long-term rate of 8%;
- that the short-rate decreases from 6% to a long-term rate of 4%.

The probabilities of each of these occurrences are 10%, 80% and 10% respectively, that is the most likely scenario is a rise in the short-rate from 6% to 8%. For each scenario we assume that the short-rate approaches the expected long-term level in exponential fashion. The expected interest-rate scenario therefore is a rise from 6% to 8%. From Figure 9.2 we see that the forward rate curve behaves differently to expected future short-rate levels. The forward rates peak at around 12–14 years, and then steadily decline as the term to maturity increases. The zero-coupon yield curve, which can be derived from the forward yield curve, has a different shape and starts to decline from the 20-year term period.

Figure 9.2 suggests that the unbiased expectations hypothesis which states that forward rates are equal to the expected level of future short-term rates, is incorrect, and so it is not valid to calculate par and zero-coupon yield curves using the expected short-rate curve. Instead the forward rate curve should be used. Figure 9.3 illustrates the extent of the error that might be made using the expected short-rate curve to calculate the zero-coupon yield curve, which is magnified over longer terms to maturity.

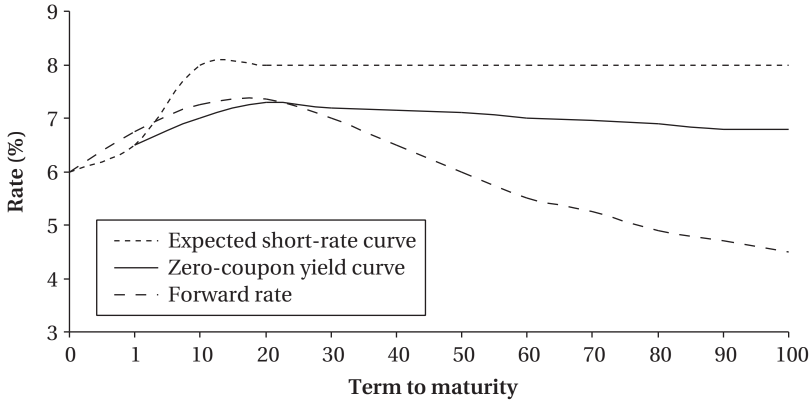


Figure 9.2: The theoretical behaviour of the long-bond yield.

From Figure 9.3 we see that to price a very long-dated bond off the yield of the 30-year government bond would lead to errors. The unbiased expectations hypothesis suggests that 100-year bond yields are essentially identical to 30-year yields, however this is in fact incorrect. The theoretical 100-year yield in fact will be approximately 20–25 basis points lower. This reflects the convexity bias in longer-dated yields. In our illustration we used a hypothetical scenario where only three possible interest-rate states were permitted. Dybvig and Marshall showed that in a more realistic environment, with forward rates calculated using a Monte Carlo simulation, similar observations would result. Therefore the observations have a practical relevance.

This is an important result for the pricing of longer-dated bonds. Certain corporate bonds including those issued by Walt Disney, Coca-Cola and British Gas to name three instances in recent years, have been very long-dated bonds, from 50 to 99 years' maturity. The analysis above suggests that such bonds priced at a spread over the 30-year government

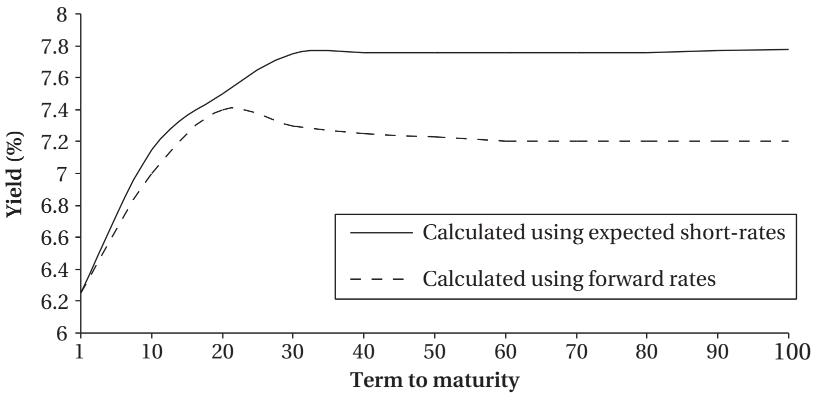


Figure 9.3: Zero-coupon yield curves calculated using expected short-rates and forward rates.

bond are theoretically undervalued. While investor sentiment would appear to demand a yield premium for buying such long-dated bonds, the theoretical credit-risk spread for a 100-year corporate bond is essentially the same as that for a 30-year corporate bond. For instance if a 30-year corporate bond has a default probability of 1% each year, while a 100-year corporate bond has a default probability of 1% for the first 30 years and a subsequent default probability of 3% for the remainder of its life, the longer-dated bond appears at first sight to hold considerably more credit risk. However it can be shown that the yield premium an investor should demand for holding the longer-dated bond will not be much more than 15–20 basis points. This is because the impact of future loss scenarios is weighted by the discount factor that applies from today to the loss date; the influence of each discount factor steadily diminishes over an increasing term to maturity.

9.2 Pricing a long bond

In a conventional positive yield curve environment it is common for the 30-year government bond to yield say 10–20 basis points above the ten-year bond. This might indicate to investors that a 100-year bond should yield approximately 20–25 basis points more than the 30-year bond. Is this accurate? As we noted in the previous section, such an assumption would not be theoretically valid. Marshall and Dybvig have shown that such a yield spread would indicate an undervaluation of the very long-dated bond, and that should such yields be available an investor, unless he or she has extreme views on future interest rates, should hold the 100-year bond.

This is intuitively apparent. In the first instance, long-dated forward rates have very little influence on the prices of bonds, and therefore for there to be a yield spread of say 20 basis points between 30 years and 40 years, forward rates would have to be very high. This reflects the relationship between spot and forward rates, the former being an average of the latter to the longest maturity. Similarly, expected future short-rates are assumed to be composed of the market's expectation of these rates and a premium for interest-rate risk. For there to be a high enough expectation such that there is a yield premium of 20 basis points between 30 and 100 years would require very high expectations about the future level of short-rates, or a very high risk premium. We now consider this in greater detail.

9.2.1 *The impact of forward rates on the long-bond yield*

From elementary financial arithmetic we know that an investment of £1 at a continuously compounded interest rate of r will have a value at time t given by e^{-rt}, so that the value of a coupon C at this time is given by Ce^{-rt}.

This enables us to set the value of a bond with a coupon of C maturing at time T and redemption value of M as (9.4):

$$\int_0^T MCe^{-rt}\mathrm{d}s + Me^{-rT} \tag{9.4}$$

which can also be given as

$$\frac{MC}{r}(1 - e^{-rT}) + Me^{-rT}. \tag{9.5}$$

We assume that the zero-coupon rate r term structure is flat until the time s and that forward rates are flat at f. The value of £1 to be received at time $t > s$ is given by:

$$e^{-rs-f(t-s)} \tag{9.6}$$

while the price of a bond maturing at T is given by (9.7):

$$P = \int_0^t MCe^{-rt}\mathrm{d}s + \int_t^T MCe^{-rs-f(t-s)}\mathrm{d}s + Me^{-rs-(T-t)f}. \tag{9.7}$$

Equation (9.7) can be integrated to give:

$$\frac{MC}{r}(1-e^{-rt}) + \frac{MC}{f}e^{-f(T-t)} + Me^{-rt-f(T-t)}. \tag{9.8}$$

This can be illustrated with an example. Consider a situation where the zero-coupon rates term structure is flat at 6% for 30 years and that forward rates are flat at f for terms from 30 years to 100 years. This results in the price of a 30-year bond with a coupon of 6% and a redemption value of £100 having a price of par, shown below:

$$P = 100 = \frac{6}{0.06}(1-e^{-0.06\times 30}) + 100e^{-0.06\times 30}.$$

Now let us imagine that the yield on a 100-year government bond with a coupon of 6.20% is 6.20%. This fits investor expectations that the very long-dated bond should have a yield premium of approximately 20 basis points. This would set the price of the 100-year bond as:

$$P = 100 = \frac{6.20}{0.06}(1-e^{-0.06\times 30}) + e^{-0.06\times 30}\left(\frac{6.20}{f}\right)(1-e^{70f}) + e^{-0.06\times 30}100e^{-f\times 70}$$

where the price of the bond is par. The forward rate given by the expression above must be greater than 6.20%, and is higher because long-dated forward rates have very little influence on the price of a coupon bond. The size of the coefficient $e^{-0.06\times 30}$, in this instance, indicates the extent of the impact of the forward rate on the price of the bond. In fact in a term structure environment that is flat or only very slightly positive out to 30 years, the zero-coupon term structure beyond this term is flat.

Let us look now at the T-period forward rate again as a function of the range of spot rates from the time t today to point T in more detail than in Section 9.1. If $P(t, T)$ is the price today of a zero-coupon bond that has a redemption value of £1 at time T, then this price is given in terms of the instantaneous structure of forward rates by (9.9):

$$P(t,\ T) = exp\left(-\int_t^T f(t,\ s)\mathrm{d}s\right) \tag{9.9}$$

where the forward rate $f(t, T)$ is given by:

$$f(t,\ T) = -\frac{\partial\ ln\ P(t,\ T)}{\partial T}. \tag{9.10}$$

However, the price of the zero-coupon bond is also given in terms of the spot rate as the expression in (9.1), where E_t is the expectation under the risk-free probability function. Therefore forward rates are related to the expected level of the instantaneous spot rates, and if we differentiate the expression in (9.10) we obtain a result that states that the forward rate is a weighted average of the range of spot rates in the period t to T. This is given in (9.11), which we encountered earlier as (9.2):

$$f(t,\ T) = \frac{exp\left(-\int_t^T r(s)\mathrm{d}s\right)}{E_t\left[exp\left(-\int_t^T r(s)\mathrm{d}s\right)\right]}. \tag{9.11}$$

In the spot-rate scenario where the expected future rate is high, the interest rate $r(T)$ will exert very little influence, while it exerts more weight at lower levels. Therefore the forward rate will be lower than the expected spot rate, and this is described below, where

$$E_t\left[r(T) \mid \int_t^T r(s)\mathrm{d}s\right]$$

is an increasing function in $\int_t^T r(s)\mathrm{d}s$.

Therefore we may write for long-term forward rates $f(t,\ T) \le E_t[r(T)]$, where interest rates are assumed to not be deterministic. This result has an important effect on the pricing of very long-dated bonds. Since forward rates lie below the level of expected short-term rates, for a very long-dated bond to trade at a yield of 6.20% means that the average level of future short-rates over the life of the bond would have to be higher than 6.20%. If this is not the case, a yield premium of 20–25 basis points between the long-bond and a very long-dated bond would indicate an unrealistically low price for the latter instrument. And crucially, for there to be a spread of this magnitude for up to say, 50 years beyond the benchmark long bond, we would observe unrealistically high forward rates and an exploding forward rate curve.

9.3 Further views on the long-dated bond yield

In the previous section we described a theorem from Dybvig, Ingersoll and Ross stating that extremely long-dated bond yields cannot decline. This carries implications about the level of interest-rate risk attached to the very long-dated yield. We present a summary of their results here.

Assume that along all random interest-rate paths ω, the short-rate gravitates towards a long-term equilibrium level of r_ω^∞, which is dependent on the path ω. A long-term level r can result if the set of interest-rate paths ω for which $r_\omega^\infty \le r$ has a positive probability. Consider then the lowest possible value r^∞ of the long-term equilibrium level r_ω^∞. The result from the previous section, that very long-dated forward rates do not reflect the unbiased expectations hypothesis but rather a disproportionate weighting of the lowest yields, implies that long-dated rates are determined by r^∞. That is, as the maturity approaches infinity, both the forward rate and the zero-coupon rate are essentially equal to the lowest possible long-term interest rate r^∞. Over time a particular long-term interest-rate level r_ω^∞ that was previously possible may become impossible, so that r^∞ may rise over time. However a previously unattainable level r_ω^∞ will remain impossible, and if it is possible today it will have been possible before. Therefore r^∞ cannot fall over time, which indicates that very long-dated forward rates and zero-coupon rates cannot fall. This means that long-dated yields are essentially given by the lowest interest-rate scenario, and will remain sticky at this level. It also means that there is a limit to the extent to which long-dated yields will be affected by changes in the expectations of future interest rates. The yield on a 100-year bond is essentially determined by the lowest yield scenario, and a fall in expected future short-term rates will have very little impact indeed.

We can illustrate this with the same example as before. Consider now that there is a 50 basis point decline in the short-rate, and that the probabilities of the three interest-rate scenarios are now:

- that there is a zero probability that the short-rate increases from 5.50% to a long-term rate of 10%;
- that there is an 80% probability that the short-rate increases from 5.50% to a long-term rate of 8%;
- that there is a 20% probability that the short-rate decreases from 5.50% to a long-term rate of 4%.

The change in the short-rate will result in a 50 basis point decline in all the expected future interest rates. However this will not result in a uniform fall in all bond yields. The impact on the zero-coupon curve and the forward rate curve is shown in Figure 9.4.

From Figure 9.4 we observe that at the very short-end, yields fall by 50 basis points. However the 100-year spot rate falls by only approximately 4 basis points, while the 100-year forward rate actually rises. This is because under the probabilities used in our scenario the 6% scenario has a higher weight at these forward dates.

What are the practical implications of these results? We may conclude that if there is a rally in the government bond market, very long-dated bond yields should be virtually unaffected. More important though, the results indicate that any term structure model that allows very long-dated yields to fall is inconsistent with the Dybvig–Ingersoll–Ross theorem, and is therefore invalid because it would permit arbitrage. Such a model would also price 100-year bonds incorrectly (although it may well price 30-year bonds correctly). The theorem is still consistent with the concept of mean reversion, and a term structure model that assumes that long-term yields will revert to a constant long-term level will fit in with

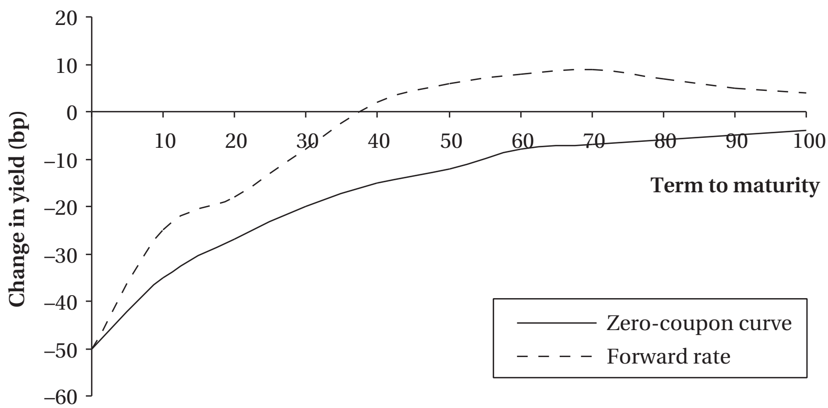


Figure 9.4: The effect of a decline in expected short rates on zero-coupon and forward curves.

the theorem. Dybvig *et al.* state that the long-term level must be the 'lowest possible level' of the average level of interest rates, but calculating this level is problematic. The issues involved in accurately pricing a very long-dated bond, and the fact that term structure modelling out to this maturity is not yet consistently applied, may explain why, despite there being no theoretical basis for it, the yields on 50-, 90- and 100-year corporate bonds sometimes lie some way above the 30-year risk-free yield.

9.4 Analysing the convexity bias in long-bond yields

In theory the results implied by a discussion of convexity imply that if there are two fixed income portfolios that have identical durations and yields, the portfolio with the higher convexity will outperform the other under conditions of a parallel yield curve shift. In fact in practice this will not be the case, as such portfolios will have lower yields, which reflects the price paid for convexity in the market.[1] We can observe this yield/convexity trade-off in the government bond yield curve: in a positive sloping yield curve environment, the yield on the longest-dated (usually 30-year) bond is almost always lower than the 15- or 20-year bond yield. This is explained by the fact that the longer-dated bond has higher convexity and that the value of this convexity is the difference in the yields. This *convexity bias* is evident in other markets, and in Chapters 41 and 43 of Choudhry (2001) we review the convexity bias that exists between the swap yield curve and the yield curve implied by long-dated interest-rate futures contracts. For example in one study[2] the ten-year interest-rate swap rate was found to be significantly lower than the rate implied by the equivalent strip of ten-year Eurodollar futures contracts. This reflects the fact that the swap instrument has convexity while the futures position does not. Therefore it is theoretically possible to benefit from a position where the trading book is short the swap (receiving fixed) and short the futures strip, as the combined effect is to be long convexity.

9.4.1 *Estimating the convexity bias*

Phoa (1998) presents an approximation of the convexity bias as follows. Consider a conventional fixed coupon bond, which has a yield at a future time t of r and a price at this time of $P(r)$. The convexity bias is estimated using

$$E[r] - r_{fwd} \approx (C/D)\sigma^2 t \tag{9.12}$$

where

$E[r] - r_{fwd}$ is the difference between the forward yield and the expected future yield (which is the convexity adjustment to the bond yield)

C is the convexity divided by two

D is the duration of the forward bond position

σ is the basis point volatility of bond yields.

The volatility value used can be estimated in two ways. We can estimate volatility separately, and then use this to calculate what the approximate convexity adjustment should be. Or we may observe the convexity bias directly and derive a volatility value from

1 For example, see Lacey and Nawalkha (1993).

2 See Burghardt and Hoskins (1995).

this. This would require an examination of market swap rates and bond yields, and use these to estimate the volatility implied by these rates.

Selected bibliography and references

Burghardt, G., Hoskins, B., 'A question of bias', *Risk*, March 1995.

Choudhry, M., *The Bond and Money Markets*, Butterworth-Heinemann, 2001.

Dhillon, U., Lasser, D., 'Term premium estimates from zero-coupon bonds: new evidence on the expectations hypothesis', *Journal of Fixed Income*, June 1998, pp. 52–58.

Dybvig, P., Ingersoll, J., Ross, S., 'Long forward and zero-coupon rates can never fall', *Journal of Business* 69(1), 1996.

Dybvig, P., Marshall, W., 'Pricing long bonds: pitfalls and opportunities', *Financial Analysts Journal*, Jan–Feb 1996, pp. 32–39.

Lacey, N., Nawalkha, S., 'Convexity, risk, and returns', *Journal of Fixed Income*, December 1993, pp. 129–145.

Phoa, W., *Advanced Fixed Income Analytics*, FJF Associates, 1998, Chapter 4.

10 The Default Risk of Corporate Bonds

In the companion volume to this book, part of the Fixed Income Markets Library, we looked at the range of corporate bond instruments that are held by investors. Institutions are interested in holding non-government bonds because of the higher yield that these bonds offer, relative to government bonds. The existence of a credit default risk on such bonds means that bondholders must ensure that the return is satisfactory and compensates them for the risk of the bond portfolio. This can be done by measuring the risk premium obtainable from the corporate bond, the total return that is expected from holding the bond, and assessing whether this is sufficient to compensate for all the risks associated with the bond, excluding the interest-rate risk. These risks must be identified and quantified, and the higher the risk, the higher the risk premium should be. A common measure for the risk premium is the *option-adjusted spread* (OAS). It is basically the spread of the corporate bond over the equivalent-maturity risk-free bond. A bond's OAS measures the constant spread that must be added to the current short-term interest rate to make the theoretical price of the corporate bond, as calculated by the pricing model, identical to the observed market price. This means that it is a quantification of the excess return of the bond over the short-rate. There is no one measure of OAS however, and it means different things depending on what type of bond it is applied to. This means that if it is used to measure the yield premium on a corporate bond that reflects a particular bond's credit risk, any specific limitations of the measure must be accounted for. There are other measures that may be considered however, in terms of the default risk of a bond, and these are considered in this chapter. We also present a theoretical default spread model.

10.1 Corporate bond default spread risk

10.1.1 *Spread risk*

The general rule of corporate bonds is that they are priced at a spread to the government yield curve. This price is a yield spread for conventional bonds or on an OAS basis for callable or other option-embedded bonds. If an OAS calculation is undertaken in a consistent framework, price changes that result in credit events will result in changes in the OAS. Therefore we can speak in terms of a sensitivity measure for the change in value of a bond or portfolio in terms of changes to a bond's OAS measure. One of these measures is the *spread duration.* The spread duration of a bond is the sensitivity of its OAS to a change in yield of one basis point. For a conventional bond, the spread duration is essentially its modified duration, because a change in the OAS would have an identical effect on the price of such a bond as a similar magnitude change in the yield on the equivalent government bond. For the same reason the spread duration of a callable bond is essentially identical to its modified duration. However the spread duration for an asset-backed security such as a mortgage-backed bond is not equal to its modified duration. This is because a change in the OAS will not have the same effect necessarily as a similar change in government bond yields.

The effect of a change in OAS for mortgage-backed bonds can be explained thus. For instance a rise in yields will lead to a rise in the level of mortgage rates, which will have the effect of decreasing prepayment rates. This will change the expected cash flow profile of the bond. However a change in the OAS of the bond will only have an effect on the bond's expected cash flows if it also leads to a rise in the prevailing mortgage interest rate.

The spread duration of a bond can be applied to calculate the break-even spread change. Remember that investors who are looking to outperform government bonds will set up portfolios to include corporate bonds, whose yields are higher than those of government bonds. However it is important for them to determine the extent to which yield spreads can widen before the additional income from the higher-yield corporate bonds is offset by the negative price effect of these bonds with regard to the price of government bonds. This measure will indicate the extent of the risk profile of their portfolios. One approach is to calculate break-even spreads for a holding period of up to one year using (10.1):

$$\text{Break-even spread} = \frac{\text{Income excess}}{\text{Spread duration}} = \frac{\text{Holding period} \times \text{spread}}{\text{Spread duration}}. \tag{10.1}$$

Example 10.1: Spread duration

An investor's corporate bond portfolio has an identical duration to a benchmark portfolio of government bonds, and an OAS of 50 basis points. Assume that the portfolio has a spread duration of five. During a 12-month holding period, the excess income of the portfolio compared to government bonds is 0.25%. How much can the OAS widen before the corporate bond portfolio begins to underperform the government portfolio?

$$\text{Break-even spread shift} = 0.25/5 = 5 \text{ basis points.}$$

Therefore if spreads widen by five basis points or more over the 12-month period, or if the OAS of the portfolio widens beyond 55 basis points, the portfolio will underperform the government portfolio.

Note that this is an approximation that is valid for short-term holdings only.

10.1.2 *Spread risk and government bond yields*

The risk premium available on a corporate bond reflects the total risk exposure of the bond, over and above the interest-rate risk which is expected to be identical in theory to the interest-rate risk on an otherwise risk-free bond. This means that the discussion of spread risk above implies that it is independent of interest-rate risk. In practice this is not so. Observation shows that the yield spread of corporate bonds is positively correlated to the outright government bond yield: when yields increase, the yield spread often decreases, while when yields fall the yield spread usually increases. Empirically though this effect can only be measured for specific issuers, and not for a class of identical credit-quality bonds. This is because the group of same-rated bond issuers is constantly fluctuating, and measuring the change in yield spread for a group of say, single-A rated borrowers will reflect changes to the group of issuers as some are re-rated, and others enter or leave the group. In most OAS calculations the relationship between outright yield levels and corporate yield spreads is not taken into account, resulting in an OAS spread of a corporate bond being equal to its nominal spread over the government yield curve.

To assess the impact of changing yield spreads therefore, it is necessary to carry out a simulation on the effect of different yield curve assumptions. For instance we may wish to analyse one-year holding period returns on a portfolio of investment-grade corporate bonds, under an assumption of widening yield spreads. This might be an analysis of the effect on portfolio returns if the yield spread for triple-B rated bonds widened by 20 basis points, in conjunction with a varying government bond yield. This requires an assessment of a different number of scenarios, in order to capture this interest-rate uncertainty.

10.2 Default risk and default spreads

10.2.1 *The theoretical default spread*

We have stated that the yield premium required on a corporate bond accounts for the default risk exposure of such a bond. The level of yield spread is determined by the expected default loss of the bond, and assumes that investors can assess the level of the default risk. This makes it possible to calculate the level of the theoretical default spread.

We set p_t as the probability that a bond will default in year t, and is the probability of default up to year t, while r_y is the expected recovery rate on the bond should it default. The default probability is assumed to fluctuate over time, while the recovery rate remains constant. Therefore the probability that the bond will not have defaulted up to the beginning of year t is given by:

$$s_t = \prod_{\tau=1}^{t-1}(1 - p_\tau) \tag{10.2}$$

while the probability that the bond will default in year t is given by:

$$p(t) = p_t s_t. \tag{10.3}$$

If the bond has a maturity of T, there are $T + 1$ scenarios, represented by default in years 1 to T or survival until maturity. The final scenario has a probability of

$$Q = s_{T+1} = 1 - \sum_{\tau=1}^{T} p(\tau). \tag{10.4}$$

Therefore using the assumed recovery rate, we may calculate the cash flows of the bond under each of the possible scenarios. For any given yield r, we can then calculate the present value of the bond's cash flows for each of these scenarios. These are denoted by $PV_1, PV_2, \ldots, PV_T$ and PV_{T+1}. Let $p(r)$ be the probability-weighted average for all the possible scenarios, shown by (10.5):

$$p(r) = \left(\sum_{t=1}^{T} p_r \cdot PV_t\right) + Q \cdot PV_{T+1}. \tag{10.5}$$

This means that $p(r)$ is the expected value of the present value of the bond's cash flows, that is, the expected yield gained by buying the bond at the price $p(r)$ and holding it to maturity is r. If our required yield is r, for example this is the yield on the equivalent-maturity government bond, then we are able to determine the coupon rate C for which $p(r)$ is equal to 100. The default-risk spread that is required for a corporate bond means that C will be greater than r. Therefore the theoretical default spread is $C - r$ basis points. If there is a zero probability of default, then the default spread is zero and $C = r$.

Generally the theoretical default spread is almost exactly proportional to the default probability, assuming a constant default probability. Generally however the default probability is not constant over time, nor do we expect it to be. In Figure 10.1 we show the theoretical default spread for triple-B rated bonds of various maturities, where the default probability rises from 0.2% to 1% over time. The longer-dated bonds therefore have a higher annual default risk and so their theoretical default spread is higher. Note that after around 20 years the expected default probability is constant at 1%, so the required yield premium is also fairly constant.

For lower-rated and non-rated bonds, the observed effect is the opposite to that of an investment-grade corporate. Over time the probability of default decreases, therefore the theoretical default spread decreases over time. This means that the spread on a long-dated bond will be lower than that of a short-dated bond, because if the issuer has not defaulted on the long-dated bond in the first few years of its existence, it will then be viewed as a lower risk credit, although the investor may well continue to earn the same yield spread.

Default probabilities are not known with certainty, and credit rating agencies suggest that higher-risk bonds have more uncertain default probabilities. The agencies publish default rates for each rating category (which are used in credit value-at-risk calculations), but the default probability values assume wider spreads for lower-rated bonds. For example a long-dated triple-B bond may have a default probability of between 0.5% and 2%, whereas a medium-dated single-B rated bond may have default probabilities of between 5% and 15%. This uncertainty will influence the calculation of the theoretical default yield spread. To estimate this, one approach involves the use of a probability distribution of the default probability, and applying the analysis using a range of possible default probabilities, rather than a single default probability. This results in a range of theoretical yield spreads. The result of this approach, somewhat surprisingly, is that the greater the range of uncertainty about the future default probability, the lower the theoretical default spread. This result has a significant impact on the yield spreads of high-risk or 'junk' bonds. The reason behind this is that an assumption of lower default probabilities results in the generation of scenarios

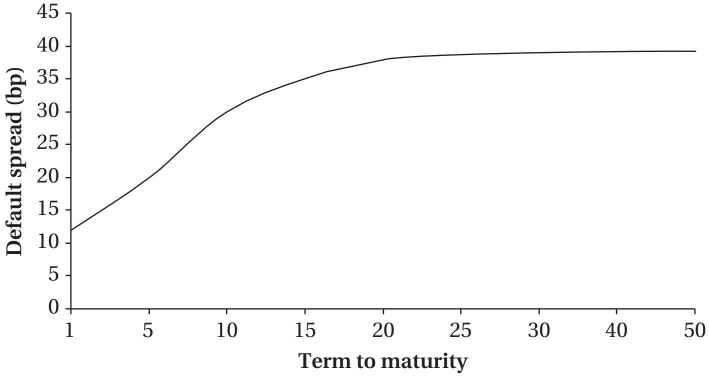


Figure 10.1: Theoretical default spread on BBB-rated corporate bond.

with higher cash flows, and the scenarios generated by these lower default probability assumptions carry a correspondingly higher weight. The default-adjusted yield being earned under a given default assumption is essentially the coupon rate minus the loss rate, where the loss rate is the product of the annual default probability and the recovery rate. Therefore the assumption of a low default probability corresponds to a lower default-adjusted yield, and has a higher weight in determining the theoretical default spread than does a high default probability assumption. A greater level of uncertainty about the level of the default probability means that more extreme high and low default probability assumptions are being used, and as the low assumptions carry greater weight in the calculation, the theoretical default spread emerges as lower.

10.2.2 *The default spread in relation to the outright government bond yield*

In the previous section we noted that in practice there is a positive correlation between the extent of the default yield spread and outright yield levels. We may wish to analyse the effect of a correlation that results in a higher level of default in a lower yield environment, or a recessionary environment when interest rates are lower. In fact the outcome will depend on whether the default probability rate that is used is assumed to rise or fall over time. If the default probability rises over time, as it does for an investment-grade bond, then the theoretical yield spread has a negative correlation with the outright yield level, whereas for a lower-rated bond or junk bond, where the default probability falls the further we move into the future, the theoretical yield spread is positively correlated with the outright yield level. This is illustrated in Figure 10.2.

Portfolio managers must also take account of a further relationship between default risk and interest-rate risk. That is, if two corporate bonds have the same duration but one bond has a higher default probability, it essentially has a 'shorter' duration because there is a greater chance that it will experience premature cash flows, in the event of default. This means that an investor who holds bonds that carry an element of default risk should in theory take this default risk into consideration when calculating the duration of his or her portfolio. In practical terms this only has an effect with unrated or junk bonds, which have default probabilities much greater than 1%. Figure 10.3 shows how the theoretical duration of a bond decreases as its assumed default probability increases.

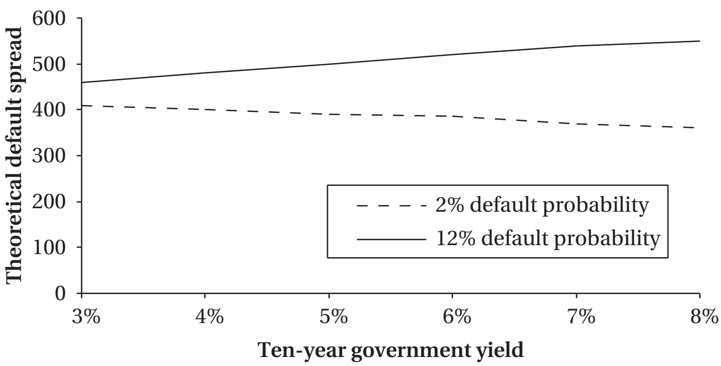


Figure 10.2: Correlation of theoretical yield spread with outright government bond yield, ten-year corporate bonds.

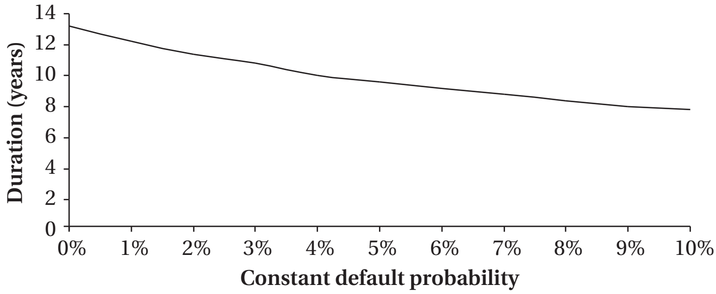


Figure 10.3: Duration of a 30-year bond relative to default probability.

Selected bibliography and references

Cohler, G., Feldman, M., Lancaster, B., 'Price of risk constant (PORC): Going beyond OAS', *Journal of Fixed Income*, March 1997.

Duffee, G., 'Treasury yields and corporate bond yield spreads: an empirical analysis', *Federal Reserve Board*, May 1996.

Duffee, G., 'Estimating the price of default risk', *Federal Reserve Board*, July 1996.

Kopprasch, R., 'Option-adjusted spread analysis: going down the wrong path?', *Financial Analysts Journal*, May–Jun 1994, pp. 121–135.

11 Brady Bonds

In this final chapter we present a discussion on Brady bonds, which are more complex fixed income instruments. Brady bonds are bonds that have been issued as part of a restructuring of a country's commercial bank loans and other debt. Existing creditors tender their loans in exchange for the new bonds, which are sovereign bonds. The name refers to Nicholas Brady, chairman of the US Federal Reserve at the time the bonds were first introduced. The Brady market is characterised by high yields and liquidity levels ranging from very liquid to illiquid. Many Brady bonds are large size issues and trade in a liquid market. Due to the features of a Brady bond, they are traded by investors taking a view on the country risk, the yield spread to Treasuries or the volatility level.

There are a number of different types of bonds in existence, for example collateralised and uncollateralised, fixed-rate and floating-rate and so on. Although most of the issues to date have been long-dated bonds, short-term Brady bonds have also been issued. The bonds are denominated in US dollars, therefore they yield a spread over Treasury bonds. This makes the bonds very interesting to trade,[1] as the yield on them reflects both the country credit risk as well as the shape and expectation of the Treasury yield curve. This is considered later in the chapter.

Countries that have completed Brady plans are listed in Table 11.1. At the end of 1998 there was over $160 billion of total debt in existence.

11.1 Brady bond structure

Brady bonds were introduced in the wake of the Latin American debt crisis of 1982. As countries threatened to default it was realised that the global banking system could be in danger if there was large-scale default. Therefore the debt was renegotiated and repacked as Brady bonds. The first step of the process is when the creditor banks negotiate with the debtor country to establish the level of borrowing that the country can realistically afford to service. This is sometimes undertaken in conjunction with the International Monetary Fund. The difference between the amount that the country can afford to service, and the actual level of its debt, is reconciled using one or more of the following:

- discount bonds, where the principal amount is cut;
- par bonds, where the debtor pays sub-market interest rates;
- a debt buy-back at an agreed discount rate.

There is an option available for creditors who do not wish to see the principal value of the loans reduced or receive below-market rates, in which case they must agree to lend additional funds in return for retaining the existing value of the obligation. Arrears in interest are usually rescheduled into a separate tradable bond.

[1] Perhaps 'interesting' is the wrong word! But they are certainly exciting products.

Argentina	Ecuador	Panama	Venezuela
Brazil	Ivory Coast	Peru	Vietnam
Bulgaria	Jordan	Philippines	
Costa Rica	Mexico	Poland	
Dominican Republic	Nigeria	Uruguay	

Table 11.1: Countries that have issued Brady bonds.

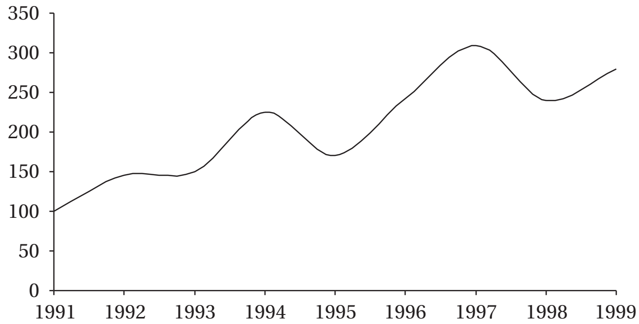


Figure 11.1: Brady bond index. Source: JP Morgan (December 1991 = 100).

Country	**Bond**	**Bid**	**Offer**
Bangladesh		73.00	78.00
Cambodia	Trade	18.00	22.00
Costa Rica	Principal A	88.25	89.25
Costa Rica	Principal B	85.25	86.25
Cuba	DEM denominated	7.00	9.00
Cuba	Yen	6.00	8.00
Guyana		20.00	30.00
Jamaica	Tranche A	90.75	92.75
Jamaica	Tranche B	86.50	87.75
Jordan	Par bonds	67.75	68.75
Jordan	Discount bonds	71.50	73.50
Jordan	IA bonds	97.00	98.00
Laos	Trade	12.00	15.00
Mongolia		10.00	15.00
Morocco	Restructured loans	92.00	93.00
Nicaragua		9.00	12.00
Sudan	Restructured loans	2.00	5.00
Surinam	Trade	23.00	27.00
Vietnam	PDI 3	53.80	54.30

Table 11.2: Exotic debt prices, US dollar denominated, February 2000. Source: IFR.

The principal amount of par and discount bonds is fully collateralised using zero-coupon US Treasury bonds. This provides comfort to creditors that the obligations will ultimately be repaid. In many cases interest payments, usually for the first few years of the issue, are also collateralised on a rolling basis; that is when the collateral backing a coupon payment is not used, it rolls forward to back the next coupon payment. This collateral arrangement therefore acts as a form of security for the investor in a Brady bond.

11.2 Types of Brady bond

The main type of Brady bond is the *collateralised fixed-rate par bond* or *Par bond*. These bonds are received in exchange for debt that is tendered at the face value. They offer the debtor permanent interest rate relief and protection from fluctuations in interest rates. The bonds are long-dated, usually of 25 or 30 years' maturity, and are conventional bullet bonds. The principal is collateralised with specially-issued zero-coupon US Treasury bonds which are held at the US Federal Reserve. There is also usually a rolling interest payment guarantee, collateralised by high-quality financial instruments rated at double-A or better, which cover 12–24 months of interest payments. In certain cases *value recovery rights* are attached to the bonds, which grant bondholders rights to payments under a formula linked to a commodity price index.

Collateralised floating-rate discount bonds or *Discount bonds* are received in exchange for eligible debt at a discount to face value, which results in permanent debt relief for the debtor in the form of partial relief on the debt obligation. In return for this the creditor receives higher interest payments, usually Libor plus 13–15%. When Mexico and Argentina issued discount bonds the yield spread was 35%, while for Bulgaria and Poland the spread was 50%. Again the bonds are long-dated, 25 or 30 years and have a single bullet payment at maturity. The principal is collateralised with zero-coupon Treasury bonds, and there is usually a rolling interest guarantee. Value recovery rights have also been attached to discount bonds.

Another type of Brady bond is the *front-loaded interest reduction bond* or *Flirb*. These bonds are received in exchange for debt at face value. They pay below-market interest rates for the first few years of their life, which is a temporary interest relief for the debtor. The principal is not collateralised, although there may be a rolling interest guarantee. As there is no collateralisation the bonds have a shorter average life, and amortise after an initial grace period of up to nine years.

Debt conversion bonds or *new money bonds* are received in exchange for debt at face value. They are conditional upon the creditor providing additional new money equivalent to a certain percentage of the amount of the eligible debt. Neither the principal nor the interest payments are collateralised, so the average lives of the bonds are shorter than par or discount bonds.

Interest arrears bonds are received in exchange for a creditor's claim on certain past due interest payments which have not been paid. There is no collateral backing for the bonds.[2] The bonds are known by a number of names, including *interest due and unpaid* or IDU bonds, *past due interest bonds* (PDI), or *interest arrears bonds* (IAB).[3]

The prices of selected Brady bonds on 5 April 2000 are given in Appendix 11.1.

2 The exception to this was in the case of bonds issued by Costa Rica.

3 In the case of Russia the bonds are known as *interest arrears notes*.

11.3 Relative value

Since Brady bonds are (in most cases) collateralised instruments, we require additional techniques when assessing their value, beyond the simple gross redemption yield measure. Essentially a par bond has three elements: the principal, which is collateralised by US Treasuries, the collateralised rolling interest guarantee, and the (risk-carrying) remaining bond cash flows. The yield on a collateralised bond will be lower than that of a non-collateralised bond because of the credit protection. The country-specific risk factor, which is in the form of the spread to the Treasury yield, should in theory apply to the risk-carrying cash flows only. This means that an investor will calculate the present value of both the principal and the collateralised interest payments and subtract this from the price of the bond. This 'stripped' price is then used when calculating the yield-to-maturity of the non-collateralised cash flows, and is known as the *stripped yield*. The stripped yield spread (as a spread to the Treasury bond) is viewed as the market's assessment of the sovereign risk. This approach enables the yields of collateralised and non-collateralised bonds to be compared. As we noted at the start of this section, the yield sensitivity of a Brady bond will reflect both the country credit risk and the Treasury yield curve. Separating a Brady bond's yield sensitivity to US interest rates and to credit risk is therefore of some importance for the market maker. We consider this here.

For a US Treasury bond we are familiar with the price/yield formula, given here as (11.1):

$$P = \frac{C/2}{(1+\frac{1}{2}r)} + \frac{C/2}{(1+\frac{1}{2}r)^2} + \frac{C/2}{(1+\frac{1}{2}r)^3} + \cdots + \frac{C/2+100}{(1+\frac{1}{2}r)^{2n}}. \tag{11.1}$$

The most common Brady bond is the par bond, for which the principal is collateralised. Given this security backing, the redemption payment can be taken as risk-free, which means that in theory it should be discounted at the Treasury yield for that maturity, rather than at a yield spread to the Treasury. Therefore the price/yield equation can be given as (11.2):

$$P = \frac{C/2}{(1+\frac{1}{2}(r+s))} + \frac{C/2}{(1+\frac{1}{2}(r+s))^2} + \cdots + \frac{C/2}{(1+\frac{1}{2}(r+s))^{2n}} + \frac{100}{(1+\frac{1}{2}rs)^{2n}} \tag{11.2}$$

where

- r is the corresponding US Treasury yield
- s is the bond credit spread over the Treasury yield
- rs is the Treasury zero-coupon yield.

The s is the stripped spread.

We apply the same analysis when calculating interest-rate sensitivity for a Brady bond. The duration D of a par bond is given by (11.3):

$$D = \frac{1}{P}\left(\begin{array}{l} \dfrac{C/2}{(1+\frac{1}{2}(r+s))}\cdot\dfrac{1}{2} + \dfrac{C/2}{(1+\frac{1}{2}(r+s))^2}\cdot\dfrac{2}{2} + \cdots \\ + \dfrac{C/2}{(1+\frac{1}{2}(r+s))^{2n}}\cdot n + \dfrac{100}{(1+\frac{1}{2}rs)^{2n}} \times n \times \dfrac{1+\frac{1}{2}(y+s)}{1+\frac{1}{2}rs} \end{array}\right). \tag{11.3}$$

The final term in the expression (11.3) receives a greater weight, because it is risk-free and therefore is discounted less heavily. This is the correct analysis; what it means is that a change in the Treasury yield curve will have a significant impact on the bond yield, as it affects all the bond's cash flows. However a change in the credit risk should have a smaller impact, because it impacts only the coupon payments. This has the effect of raising the duration of a par bond, compared to the duration calculation carried out using the conventional approach. Using this analysis, the yield sensitivity of a par bond, with respect to changes in the Treasury yield, is given by (11.4):

$$\frac{\Delta P/P}{\Delta y} = -\frac{D}{1+\frac{1}{2}(r+s)}. \tag{11.4}$$

The credit risk sensitivity is a function of changes in the credit quality of the bond, and so it is not symmetrical with the bond's Treasury yield sensitivity. This is because the credit risk yield premium is only applicable to the bond's coupon payments, so the sensitivity measure will be lower. It is given by (11.5):

$$\frac{\Delta P/P}{\Delta s} = -\frac{D-(A/P)}{1+\frac{1}{2}r+s} \tag{11.5}$$

where $A = n \cdot 100 \cdot \dfrac{1+\frac{1}{2}(r+s)}{(1+\frac{1}{2}rs)^{2n+1}}$.

The presence of the A/P term has the effect of reducing the impact of a change in the yield premium on the price of the bond compared to a change in the Treasury yield.

Appendix

Appendix 11.1: Brady bond prices

	Bond	Price		Bond	Price
Argentina	Par	69.78	Mexico	Discount A	97.94
	Discount	84.13		Par A	83.41
	FRB	91.33		Par B	84.88
Brazil	IDU	99.59	Nigeria	Par	71.70
	'C'	72.00	Panama	PDI	83.73
	Par	65.17		IRB	79.57
	Discount	76.54	Peru	PDI	65.85
	EI	88.69		FLRB	60.73
	DCB	71.79	Philippines	FLB B	95.29
	NM94	83.98		Par B	85.08
	EXIT	68.92	Poland	Discount	99.75
	FLRB	76.17		RSTA	66.67
Bulgaria	IAB	78.10		Par	62.17
	Discount A	80.10	Russia	IAN	27.88
	FLB A	72.25		PRIN	25.51
Croatia	FRN A	92.84	Venezuela	FLB A	79.17
	FRN B	84.88		FLB B	79.17
Ecuador	PDI	26.29		DCB	77.88
	Discount	41.07		Discount A	76.42
	Par	36.67			
Jordan	Discount	78.18			
	Par	60.67			

Table 11.3: Brady bond prices, 5 April 2000. Source: Bloomberg.

Selected bibliography and references

Alumian, A., *The Emerging Bond Markets*, Mars Publishing, 1998.

Brauer, J., 'Brady bonds', in Fabozzi, F. (ed.), *The Handbook of Fixed Income Securities*, 5th edition, Dow-Jones Irwin, 1997.

Kamin, S., von Kleist, K., 'The evolution and determinants of emerging market credit spreads in the 1990s', *Bank for International Settlements Working Paper*, No. 68, May 1999.

The Economist, 'Another coup', *The Economist*, 27 November 1999.

Index

Index prepared by Terry Halliday

Adams, K.J., 118
Adapted filtration, concepts, 48
Affine term structure models:
  see also Equilibrium . . . ; Multi-factor . . . ;
    One-factor . . .
  concepts, 56–7, 68
Anderson, N., 97, 99, 102–4, 108–13, 118, 136–7
Arbitrage, xiv, xvii, 5, 49, 55, 60–1, 68–75, 82–3, 90–3, 160–1
  arbitrage-free models, xiv, xvii, 49, 55, 68–75, 82–3, 87, 90–3
  futures trading, 5
  long bonds, 160–1
  realistic approaches, 90
Argentina, 170, 171, 173
Arithmetic averages, uses, 58
Asset prices, xiii–xiv, 35–56, 60–76, 82, 88–9
  behaviour, 35–43
  discrete-time environment, 42–3
  dynamics, 35–54, 88–91
  generalised Weiner processes, 40–2
  information structures, 48–50
  Itô process, xxi, 42, 44–50, 51–3, 75, 81–2
  jumps, 35, 37, 87–8
  martingale property, 39–40, 49–50, 81–4
  model derivation, 42–54
  perfect markets, 48–50
  Weiner processes, xvi–xvii, xx, 37–44, 51–2, 67, 80–3, 85–6
Attari, M., 88
Augmented filtration, concepts, 48
Australia, 141–2, 144
Average cost curves, 59

B-splines, *see* Basis splines
Bank of England, 98, 102–4, 108, 110–12, 143–4, 146–50
Banks:
  see also Central banks
  base-rate changes, 88
  interest-rate models, 69, 88
Barone, E., 113
Basis point value (BPV), 11–12, 14–15, 20–3, 28–31
Basis splines, 102, 109
Basis spread trades, 3–5
Baxter, M., 35, 38–40, 79, 80, 83, 85
Baz, J., 88
BDT, *see* Black–Derman–Toy model
Beim, D., 117
Benchmark uses, government bonds, 1, 8, 10–11, 21, 23, 27, 91
BGM model, 87
Binomial lattice approach:
  Black–Derman–Toy model, 73
  Ho and Lee model, 69–71
Bjork, T., 88, 104, 109, 117
Black–Derman–Toy model (BDT), xvii, 61, 73, 89–90, 92
Black–Scholes model, xiii, xiv, xix, 49–50, 88–9
Bliss, R., 97
Bloomberg, 23–6, 31–3, 145, 173
Bond markets, information, 99–102
The Bond and Money Markets (author), 2, 146
Bond prices, 46–7, 55–6, 60–76, 82, 84–9, 97–102, 154–61
  see also Asset prices
  equation, 46–7, 100
  spot/forward rates, 60–1, 73–5, 97–102, 157–8
  zero-coupon bonds, 69–73, 75, 84–7, 154–9
Bond spreads:
  see also Spread . . .
  types, 12–13
  weighting, 11–12, 20–3, 27–8
Bootstrapping techniques, 119
Boundary condition, cubic splines, 126
BPV, *see* Basis point value
Brace, A., 87
Brady bonds, 169–74
  background, 169–74
  concepts, 169–71
  issuing countries, 169–70, 173
  relative values, 171–3
  stripped yield, 172
  structure, 169–71
  types, 169, 171
  value recovery rights, 171

Brazil, 170, 173
Brennan, M., 57, 84–5
Brennan–Schwartz model, 84–5
British Gas, 156
Briys, E., 44
Brownian processes, xvi–xvii, xx, 37–53, 61, 81–3, 85–6
  see also Weiner processes
Bulgaria, 170, 171, 173
Bull markets, 27
Burden, R., 118, 130
Butterfly spreads, concepts, 5–6, 13, 27–33
Butterfly/barbell trades, 13, 27–33
  background, 13, 27–33
  Bloomberg 'BBA' analysis, 31–3
  concepts, 13, 27–33
  convexity, 30–2
  portfolio yield, 28, 29–30
  putting on the trade, 28–9
  yield curves, 27–33
  yields, 27–33, 28
Buying the spread, 5

C++, 132
Calibration processes, interest-rate models, xiv–xv, 69–75, 79, 88–9, 92–3, 97–115
Callable bonds, spread duration, 163–4
Campbell, J., 59
Canada, 141
Case studies, hedging, 16–18
Central banks:
  see also Banks; Government authorities
  base-rate changes, 88
  fitted curves, 97–9, 102–4
  forward interest rates, 98–104
Chan, K., 64
Charting, *see* Technical analysis
Christensen, B., 104, 109, 117
CIR, *see* Cox–Ingersoll–Ross model
Clamped splines, boundary condition, 126
Coca-Cola, 156
Collateralised fixed-rate par bonds, *see* Par bonds
Collateralised floating-rate discount bonds, *see* Discount bonds
Complete markets, 48–50
Consol models, 84–5
Continuous random variables, concepts, xviii–xix, 35–7, 44, 47–53
Continuous time, spot/forward rate relationship, 57–61, 70–1
Convexity, xiii, xiv, 16–18, 27, 153, 155, 161–2
  bias estimates, 161–2
  butterfly/barbell trades, 30–2
  long bonds, 153, 155, 161–2
  portfolio management, 161–2
Cooper, I., 85
Corporate bonds, 1–2, 98–100, 163–8
  credit ratings, 164–8
  default risks, 163–8
  government bonds, 163–8
  institutional investors, 163
  long bonds, 156–7
  OAS, 163–5
  risks, 163–8
  spread risk, 163–8
Corporate Bonds and Structured Financial Products (author), xiv, xv, 8, 20
Corporate finance, 150
Correlations, hedge analysis, 14–15
Costa Rica, 170, 171
Costs:
  average cost curves, 59
  of carry, 5
  marginal cost curves, 59
  relative value trading, 5, 23
Country-specific risks, 172–3
Coupon, 7–8, 10, 19, 24–7, 91, 100, 110–11
  spreads, 24–7
  yield curves, 7–8, 10, 19, 24–7
  yield relationship, 10, 19
Cox–Ingersoll–Ross model (CIR), xiii–xiv, xvii, 49, 56–7, 61–2, 68, 75, 92, 98
Credit default risks, xiv, 8–9, 19–20, 99, 163–8, 172–3
  see also Risks
  concepts, 163–8
  corporate bonds, 163–8
  duration, 163–5, 167–8
  expected values, 165–8
  government bonds, xiv, 8–9, 19–20, 99
  measures, 163–8
  yield determinants, 19–20
Credit ratings, 164–8, 173
Cuba, 170
Cubic spline application (CUBED), 132
Cubic splines, 97, 101–2, 104–15, 117–39, 147–8
  Anderson–Sleath approach, 108–13, 118, 136–7
  boundary condition, 126
  concepts, 97, 101–2, 104–15, 117–39, 147–8
  criticisms, 101, 117, 135

empirical proof, 132–4
errors, 133–4
flexibility benefits, 119
forward rate curves, 135–7
Gaussian elimination, 125, 128–32, 137–8
iterative solution, 125, 127–8, 130, 132
mathematics, 118–32
methodology, 119–37
natural/clamped splines, 126
practical approach, 121–32
software, 128, 132
supposition, 120–1
table of related formulas, 124–5
tri-diagonal matrix, 125, 130–2, 137–8
Waggoner model, 105–9, 118, 136–7
Curvature trades, 13

Dahlquist, M., 97
Das, S., 88
de la Grandville, O., 118
Deacon, M., 113, 147, 148
Debt buy-backs, 169
Debt conversion bonds, 171
Default risks, *see* Credit default risks
Default yield spread, 165–8
Derivatives, xiv, xv–xviii, 1, 35, 44–6, 74, 82, 88–91, 99–102, 161, 163–5
see also Futures; Options
hedging, 2–4, 14
liquidity issues, 2–3
prices, 35, 44–6, 48–50, 56–7, 82, 88–91, 99–102
probability measures, 82–3
speculative trading, 2–4
Derry, A., 113, 147, 148
Deterministic models, 36, 159
Deutsche Bank, 117
Diffusion coefficient, 45, 49
see also Volatility
Discount bonds, 169, 171–2
Discount function, 60–1, 69–71, 89, 91, 98–102, 104–15, 118–38, 146–51, 154–7
Discrete random variables, xviii, 42–3, 58
Discrete-time environment, asset prices, 42–3
Distribution, risk-free interest rates, 43
Dollar value of a basis point (DVBP), 11–12, 20
see also Basis point value
Drift components:
processes, 36, 38, 42–3, 45, 51–2, 69, 73, 81–4, 93
short-rate models, xvi–xvii, 61–4, 69, 73

Duffie, D., 37, 56, 113
Duration, xv, 5–12, 16–18, 20–3, 28, 30, 142–5, 163–8
credit default risks, 163–5, 167–8
index-linked bonds, 142–5
modified duration, xv, 11–12, 21, 28, 30, 163–4
spread duration, 163–5
Duration-weighting method, 23, 27–8
DVBP, *see* Dollar value of a basis point
Dybvig, P., 153, 155, 156–7, 159–61
Dynamic adjustments, portfolios, 23

Economic factors, xv, 141–2
The Economist, 1
Edmister, R.O., 102
Efficient markets, xx, 48–50
Equilibrium models, xvii, 55–7, 63–8, 75, 79, 91–3
Estimation methods, 92, 94–5, 97–116, 142
Eurobonds, hedging examples, 16–18
European-style contingent liability, valuation, 50
Excel, 128
Expectations hypothesis, 19–20
Expected values:
concepts, xix, 165–8
credit default risks, 165–8
Exponential functions, 36
Extended Vasicek model, *see* Hull–White model

Fabozzi, F., 29, 117
Faires, D., 118, 130
Fama, E., 48
Federal Reserve, 98
Federal Reserve Bank of Atlanta, 105
Filtration concepts, 48–9
First-order risk-neutral position, 19–22, 27
Fisher identity, 146, 149
Fisher, M., 99, 104–9, 118, 136
Fitting processes:
interest-rate models, 69–75, 79, 88–9, 92–3, 97–115
yield curves, 97–115, 117–39, 146–7
Flirbs, *see* Front-loaded interest reduction bonds
Fong, H.G., 101–2, 117
Foresi, S., 88
Forward interest rates, 3–5, 56–77, 79–93, 97–115, 135–7, 154–7, 160–1
bond prices, 60–1, 97–102
central banks, 98–104

continuous time, 57–61, 70–1
definition, 58, 80–1, 97–8
Heath–Jarrow–Morton model, 79–92
implied rates, 98–102, 146–9
long bonds, 154–9
spot rates, 55–61, 73–6, 97–102, 110–11, 135–6, 157–8
Forward rate curves, 97–115, 135–7, 155–7, 160–1
Front month contracts, 3
Front-loaded interest reduction bonds, 171
Fundamental analysis, 2
Futures, 1–6, 14, 97, 101, 161
see also Derivatives
background, 2–6
designatory letters, 3
trading, 2–6

Gatarek, D., 87
Gaussian elimination, cubic splines, 125, 128–32, 137–8
Gaussian processes, xvi, xviii–xix, 44, 47–50, 56–7, 61–3, 66–8, 71–5, 85, 125, 137–8
Gaussian uncertainty, 48–50
GCV, *see* Generalised cross-validation
General Hull–White model, *see* Hull–White model
General method of moments, 92
Generalised cross-validation (GCV), 107
Generalised Itô formula, 47–8
see also Itô . . .
Generalised Weiner processes, 40–2
Geometric averages, uses, 58, 80, 102–3
Geometric Brownian motion, 42–3
see also Brownian processes
Gilts, xiv, 1–2, 5–6, 6–8, 19–33, 99–102, 163–5
see also Government bonds
hedge ratio, 5
pensions funds, 8
price spreads, 1–2
relative value trading, 8–11, 19–33
yield curves, 1, 5, 6–8, 103–5
'Goal Seek' function, Microsoft Excel, 128
Government authorities:
see also Central banks
fitted curves, 97–9, 102–4, 112–13
repo market, 112–13
yield curves, 7–8, 97–9, 112–13
Government bonds, xiv, xv, 1, 6–11, 99–102, 163–5
see also Gilts
benchmark uses, 1, 8, 10–11, 21, 23, 27, 91, 99
corporate bonds, 163–8
credit default risks, xiv, 8–9, 19–20, 99
index-linked bonds, xiv, xv, 141–51
relative value trading, 8–11, 19–33
spread risk, 164–8
yield determinants, 6–8

Hardman, Ed, 4
Harrison, J., 40
Harrison, M., 82
Heath, D., 80
Heath–Jarrow–Morton model (HJM), xiv, 55, 56–7, 79–92, 97–8
background, xiv, 55, 56–7, 79–92, 97–8
concepts, 80–4, 85–93, 97
multi-factor model, 85–93, 98
no-arbitrage principles, 83–4
single-factor model, 80–4
Hedge analysis, 14–15
Hedging:
approaches, 2–4, 5, 13–18
background, 2–4, 5, 13–18
case studies, 16–18
concepts, 2–4, 5, 13–18
correlations, 14–15
derivatives, 2–4, 14
Eurobonds/US Treasuries, 16–18
gilts, 5
index-linked bonds, 144
optimum hedge equation, 15, 18
simple approaches, 14
volatility, 14–15
Historical analysis, yield curves, 12–13
HJM, *see* Heath–Jarrow–Morton model
Ho and Lee model, 69–71, 74–5, 79, 89–90
Honore, P., 88
Horizon pricing, 91
Hull, J., 44, 79
Hull–White model, xiii–xiv, 56–7, 71–3, 74–5, 89–90, 92, 97, 109

IABs, *see* Interest arrears bonds
IDU, *see* Interest due and unpaid bonds
Imai, K., 55
IMF, *see* International Monetary Fund
Implied forward rates, 98–102, 146–9
In-the-money options, 89

Index-linked bonds, xiv, xv, 141–51
  see also Inflation rates
  advanced analytics, 141–51
  background, xiv, xv, 141–51
  concepts, 141–51
  duration, 142–5
  hedging, 144
  institutional investors, 143
  pensions funds, 143
  portfolio management, 143–5
  prices, 141–51
  real yields, 141–5, 149–51
  risks, 143–5
  taxation, 141, 145, 148, 150
  term structure of interest rates, 146–7, 149–50
  time-lags, 141, 144–5
  volatility, 141
Inflation rates, xv, 19–20, 62, 98–9, 141–51
  see also Index-linked bonds
  estimation methods, 142, 146–51
  expectations, 142, 146–51
  information, 142, 146–51
Information:
  asset prices, 48–50
  bond markets, 99–102
  inflation expectations, 142, 146–51
  yield curves, 98–102
Ingersoll, J., 35, 153, 159–61
Initial probability measures, 80–1
Institutional investors:
  corporate bonds, 163
  index-linked bonds, 143
  liquidity issues, 8, 20
Insurance investment management, 150
Integrals, 39, 47, 53–4, 56, 71–3
Interest arrears bonds (IABs), 171
Interest due and unpaid bonds (IDU), 171
Interest rates:
  see also Forward ...; Spot ...; Term structure ...
  base-rate changes, 88
  high-coupon bonds, 27
  jumps, 35, 37, 87–8
  long bonds, 154–61
  market direction, 6–9, 11
  negative forward interest rates, 49, 62–3, 67–8, 75
  processes, 61–3
  rapid-growth economies, 141–2
  risk-free interest rates, xvii–xviii, 43, 63–5, 80–1, 84, 92, 110–11
  risks, 49–50, 165–8
  uncertainties, 49–50
  volatility, 61–3, 68–74, 79–93
  yield, 57–61, 63–4
Interest-rate models, xiii, xiv–xvi, 35, 55–77, 79–95
  arbitrage-free models, xiv, xvii, 49, 55, 68–75, 82–3, 87, 90–3
  background, xiii, xiv–xvi, 35, 55–77, 79–95
  BGM model, 87
  Black–Derman–Toy model, xvii, 61, 73, 89–90, 92
  Brennan–Schwartz model, 84–5
  calibration processes, xiv–xv, 69–75, 79, 88–9, 92–3, 97–115
  concepts, 56–77, 79–95
  Cox–Ingersoll–Ross model, xiii–xiv, xvii, 49, 56–7, 61–2, 68, 75, 92, 98
  differences, xiv–xv
  equilibrium models, xvii, 55–7, 63–8, 75, 79, 91–3
  fitting processes, 69–75, 79, 88–9, 92–3, 97–115
  Heath–Jarrow–Morton model, xiv, 55, 56–7, 79–92, 97–8
  Ho and Lee model, 69–71, 74–5, 79, 89–90
  Hull–White model, xiii–xiv, 56–7, 71–3, 74–5, 89–90, 92, 109
  Merton model, 49, 67, 75
  model selection, 89–91
  multi-factor models, xiii–xiv, xv–xvi, 55–7, 79, 84–93, 98, 102
  one-factor models, xiii–xiv, xv–xvi, 55–75, 79, 84–5, 88–93, 98, 102–3
  realistic approaches, 90–1
  relative value trading, 92–3
  two-factor models, xv–xvi, 56, 84–5
  Vasicek model, xiii–xiv, xvii, 55–6, 61, 64–8, 71, 75, 88, 92, 97, 109
  whole-yield curve models, 55, 56–7, 79, 97
Interest-rate quanto options, 89
International Monetary Fund (IMF), 169
Inverted yield curves, 1, 7–8, 19–20, 23, 69
Investors, expectations, xv, 7–8, 98, 110–11
Iterative solution, cubic splines, 125, 127–8, 130, 132
Itô process, xxi, 42, 44–53, 75, 81–3

James, J., 35, 40, 55, 56–7, 71, 79, 84–5, 97, 101–2, 109, 117
Jamshidian model, 57, 87
Japan, 63

Jarrow, R., 55, 56, 80, 84, 88
Jump-augmented HJM model, 88
Jump-augmented Vasicek model, 88
Jumps:
 asset prices, 35, 37, 87–8
 models, 87–8
'Junk' bonds, 166–7

Kernel approximations, 97
Kim, J., 97
Knot points, polynomial segments, 101–2, 104–15, 118–38
Kreps, D., 40

Lagrange polynomial, 118
Laguerre function, 102–3
Lancaster, P., 115
Latin American debt crisis, 1982, 169–71
Leverage components, futures trading, 2–6
Lévy theorem, 44
LIBOR market models, 87, 97
LIFFE, 2–6
Linear interpolation, cubic splines, 119, 135–7
Linear programming techniques, 97
Liquidity issues:
 Brady bonds, 169–74
 derivatives, 2–3
 institutional investors, 8, 20
 yield curves, 7–8, 13
 yield determinants, 6–8, 19–20, 91
Lo, A., 59
Local part of the yield curve, 7–8, 9–10, 23–4, 112–13
Logarithmic functions, 36
Lognormal distributions, xix, 43, 46, 61–3, 73
Long bonds, 3–5, 13–18, 20–2, 27–8, 109–10, 153–62
 convexity, 153, 155, 161–2
 corporate bonds, 156–7
 forward interest rates, 154–9
 interest rates, 154–61
 prices, 157–61
 returns, 153–4
 theories, 153–7, 159–60
 yield analysis, 153–62
 zero-coupon yields, 153–4, 159–60
Long doubles, C++ variables, 132
Long-dated forward rates, 154–7
Long-dated options, 89
Long-dated yields, 154–7

Macaulay duration, 29
McCulloch method, 101–2, 105–6, 109, 113–14, 117–18, 146–7, 148
MacKinley, A., 59
Madan, D., 88, 102
Marginal cost curves, 59
Marginal rates, 59
 see also Forward interest rates
Market makers, 1–4, 6–7
Markets:
 bull markets, 27
 information, 99–102
 perfect markets, xx, 48–50
 repo markets, 12, 20, 23, 25, 63, 99, 112–13
 transparency, 1–2
Markov processes, xvi, xx
Marshall, W., 153, 155, 156–7
Martingale property, 39–40, 49–50, 81–4
Mathematics requirements, xviii–xxi, 118–32
Matrices, tri-diagonal matrix, 125, 130–2, 137–8
Matthews, Peter, 4
Maximum-likelihood estimations, 92
Mean, 36–43, 64–7, 71–3
Mean reversion, 36–7, 39, 57, 64–7, 71–3, 75, 93
Merton, R., 49, 67, 75
Mexico, 170, 171, 173
MFR, *see* Minimum Funding Requirement
Microsoft Excel, 128
Minimum Funding Requirement (MFR), 8
Modified duration, xv, 11–12, 21, 28, 30, 163–4
Money market account processes, spot rates, 81–2, 87, 101
Money market yield curves, independent movements, 23–4
Monte Carlo simulation, 156
Mortgage-backed bonds, 89, 150–1, 163–4
Morton, A., 80
Multi-factor models, xiii–xiv, xv–xvi, 55–7, 79, 84–93, 98, 102
 see also Heath–Jarrow–Morton model
 assessment, 88–93
 concepts, 84–93
 model selection, 89–91
Musiela, M., 87

Natural splines, boundary condition, 126
Neftci, S., 35–6, 37, 39, 44, 47, 79
Negative forward interest rates, 49, 62–3, 67–8, 75

Nelson, C., 102–4, 109, 114–15, 117
Netherlands, 141
New money bonds, 171
New Zealand, 141
Newtonian terms, 39
Nigeria, 170, 173
No-arbitrage principles, 83–4, 90–1, 153–4
Noise, 36, 38–9, 45–6
  see also Volatility
Nominal interest rates, components, 62
Non-economic factors, xv
Normal distributions:
  see also Gaussian . . .
  concepts, xviii–xix, 44, 47, 52, 61–3, 66–7, 71–3
Nychka, D., 99, 104–9, 118, 136

OAS, *see* Option-adjusted spread
One-factor models, xiii–xiv, xv–xvi, 55–75, 79, 84–5, 98, 102–3
  assessment, 88–93
  concepts, 55–6, 63–8, 79, 84–5
  model selection, 89–91
Open position trading, 2–5
Optimum hedge equation, 15, 18
Option-adjusted spread (OAS), 163–5
Option-embedded bonds, 163–5
Options, xv–xviii, 2, 35, 44, 74, 88–91, 163–5
  see also Derivatives
  prices, 35, 44–6, 48–50, 56, 88–91
  volatility, 74
Ornstein–Uhlenbeck process, 39, 64–7
OTC options, 74
Out-of-the-money options, 89
Overview, xiii–xiv

Par bonds, 169, 171–3
Par yield curves, 8–11, 98
Parametric techniques:
  concepts, 102–3, 108, 109–10, 114–15, 117–18, 136
  curve-fitting techniques, 102–3, 108, 109–10, 114–15, 117–18, 136
  drawbacks, 136
Parsimonious models, *see* Parametric . . .
Past due interest bonds (PDI), 171
Path, *see* Realisation
PDI, *see* Past due interest bonds
Pelsser, A., 71
Pension funds:
  gilts, 8
  index-linked bonds, 143
Perfect markets, xx, 48–50
Peru, 170, 173
Phoa, W., 61, 153, 154, 161
Piecewise polynomials, concepts, 104–5, 118–38
Pienaar, R., 117
Pliska, S., 40, 82
Poisson processes, 37–9, 88
Poland, 170, 171, 173
Polynomial approximations, yield curves, 101–2, 104–5, 118–38
Portfolio management, xiv–xv, 16–18, 19, 23, 27, 143–5, 164, 167–8
  butterfly/barbell trades, 28, 29–30
  convexity, 161–2
  credit default risks, 164, 167–8
  index-linked bonds, 143–5
  spread duration, 164
  weightings, 29–30
  yield, 29–30
Present value of a basis point, *see* Basis point value
Probability:
  density functions, xvi, 36
  derivatives, 82–3
  distributions concepts, xviii–xix
  expected values, xix, 165–8
  initial probability measures, 80–1
Putting on the trade, butterfly/barbell trades, 28–9

Quanto options, 89
Quartic splines, 118

'Ragged' curves, 101
Random variables, concepts, xviii–xix, 35–7, 44, 47–53, 159–60
Random walk, 44, 47
Real interest rates, xv
Real yields, index-linked bonds, 141–5, 149–51
Realisation, random processes, xx
Realistic approaches, model selection, 90–1
Rebonato, R., 79, 85
Regression analysis, concepts, xix–xx, 105
Regression cubic splines, 105–7
Relative value trading, xiii, 1–5, 6–13, 19–33, 92–3
  background, 1–5, 6–13, 19–33
  butterfly/barbell trades, 13, 27–33
  concepts, 1–5, 11–13, 92–3
  costs, 5, 23

coupon spreads, 24–7
government bonds, 8–11, 19–33
interest-rate models, 92–3
risk weighting, 20–3, 27–8
yield curves, 6–11, 19–33, 92–3
yield spread trades, 23–5, 163–5
Rennie, A., 35, 38–40, 79, 80, 83, 85
The Repo Handbook (author), 23
Repo markets, 12, 20, 23, 25, 63, 99, 112–13
Returns:
long bonds, 153–4
risks, 163
Risa, S., 113
Risk-free interest rates, distribution, 43
Risk-free zero-coupon rate, xvii–xviii, 63–5, 80–1, 84, 92, 110–11
see also Short...
Risk-neutral approaches, xvii–xviii, 5, 19–20, 27, 49–50, 90–1
Risks:
see also Credit default...
Black–Scholes model, 49–50, 88–9
corporate bonds, 163–8
country-specific risks, 172–3
index-linked bonds, 143–5
market direction, 6–9, 11
OAS, 163–5
premiums, 69–71, 92, 157–8, 163–8, 172–3
returns, 163
spread-trade weighting, 20–3, 27–8
yield, 19–20
Ross, S., 35, 49, 153, 159–61
'Roughness penalty', 105–9, 136, 147
Russia, 171, 173

Salkauskas, K., 115
Sample space, xviii
Scaling concepts, 22
Schaefer, S., 147
Schwartz, E., 57, 84–5
SDEs, *see* Stochastic differential equations
Selling the butterfly spread, 6
Shea, G., 102
Short-term rates, xvi–xviii, 2–8, 13–18, 20–2, 27–8, 49–50, 55–6, 61–77, 98–9, 155–61
Short-term repo, fitted curves, 113
Shorting the spread, 5
Siegel, A., 102–4, 109, 114–15, 117
Simultaneous equations, 118
Single-factor HJM model, 80–4
see also Heath–Jarrow–Morton model
Sleath, J., 99, 102–4, 108–13, 118, 136–7
Smoothing techniques, 92–3, 106–7, 114–15, 117–38
Software, cubic splines, 128, 132
Sovereign debt, Brady bonds, 169–74
Specialness extent, repo rates, 20, 23, 63, 113
Speculative trading, concepts, 2–4
Spline techniques, 97, 101–15, 117–39, 147–8
Spot rate curves, 97–9, 104–5, 109–12, 135–6, 155–6
Spot rates:
see also Interest rates
bond prices, 60–1, 84, 97–102, 157–8
concepts, 58–61, 73–6, 80, 97–102, 104–5, 109–12, 135–6
definition, 58, 80
forward interest rates, 55–61, 73–6, 97–102, 110–11, 135–6, 157–8
money market account processes, 81–2, 87, 101
Spread duration, 163–5
Spread risk:
corporate bonds, 163–8
government bonds, 164–8
Spread trading:
see also Relative value trading
concepts, 2, 11–13, 19–33
risk weighting, 20–3, 27–8
spread types, 12–13
Square-root processes, 56, 61–3
Standard deviation, 38–43, 51–2, 63–7, 69–75
State variables, 92
Steeley model, 56
Stochastic calculus, concepts, xx–xxi, 39, 44–53
Stochastic differential equations (SDEs), 45, 53, 82–3
Stochastic integrals, 47, 53–4, 56
Stochastic processes, xx–xxi, 35–53, 64–7, 80
concepts, xx–xxi, 35–53, 80
information structures, 48
Stripped yield, Brady bonds, 172
Strips, 3, 111, 161
'Stub', 3
Submartingale processes, 40
Suits, D., 104
Supermartingale processes, 40
Svensson, L., 97, 99, 102–4, 109–12, 114–15, 117
Swaps, 97, 101, 161–2
Sweden, 103, 141
Synthetic bonds, 99

T-bills, *see* Treasury bills
Taxation, 19, 91, 141, 145, 148, 150
Taylor expansion, 45, 51–2
Technical analysis, 2
Tenor points, cubic splines, 118–38
Term structure of interest rates, xiii–xv, 6–8, 19–20, 49–50, 55–76, 79–93, 117–39, 146–51
  estimates, 97–116, 117–39, 146–51
  index-linked bonds, 146–7, 149–50
  multi-factor models, xiii–xiv, xv–xvi, 55–7, 79, 84–93, 98, 102
  spot/forward rates, 61, 97–9
  volatility, 79–93
Term-to-maturity, 6–8, 19–23, 58–9, 64–7
Theoretical default spread, 165–8
Tick prices, 19
Time-series data, 68–9
TIPS, *see* Treasury inflation-indexed securities
Trading:
  approaches, 1–18
  Brady bonds, 169–74
  concepts, 1–2
  futures, 2–6
  relative value trading, xiii, 1–5, 6–13, 19–33, 92–3
Transparency, markets, 1–2
Treasury bills (T-bills), repo market, 112–13
Treasury inflation-indexed securities (TIPS), 141, 145, 148
Tri-diagonal matrix, concepts, 125, 130–2, 137–8
Turnbull, S., 84
Two-factor models, xv–xvi, 56, 84–5

UK, *see* United Kingdom
Uncertainties:
  Gaussian uncertainty, 48–50
  interest rates, 49–50
Uncountable numbers, 39
United Kingdom (UK):
  see also Gilts
  benchmark uses, 98
  fitted curves, 97–9, 102–4
  forward interest rates, 98, 135
  index-linked bonds, 141, 143–6, 148–51
  inflation expectations, 146
  long bonds, 153
  transparency, 1–2
  zero-coupon bond market, 110–11
United States (US):
  Brady bonds, 169–74
  coupon spreads, 24–7
  forward interest rates, 98
  hedging examples, 16–18
  index-linked bonds, 141, 144–5
  long bonds, 153
  transparency, 1–2

Value recovery rights, Brady bonds, 171
van Deventer, D.R., 118
van Deventer, J., 55
'Variable roughness penalty' (VRP), 105, 108–9, 136
Variance, xix, 38–43, 51–2
Vasicek model, xiii–xiv, xvii, 55–6, 61, 64–8, 71, 75, 88, 92, 97, 109
Vasicek, O., 49, 64, 97, 101–2, 117
Volatility, xvi–xvii, 36, 38–9, 45–6, 61–3, 68–74, 79–81, 93, 161–2
  futures trading, 2–6
  hedging approaches, 14–15
  index-linked bonds, 141
  interest rates, 61–3, 68–74, 79–93
  options, 74
  short-rate models, xvi–xvii, 61–2, 68–73
  term structure of interest rates, 79–93
Volatility smiles, 89
VRP, *see* 'Variable roughness penalty'

Waggoner, D., 97, 99, 102, 104–9, 118, 136–7, 146–7, 150
Walt Disney, 156
Web site, author, 81
Webber, N., 35, 40, 55, 56–7, 71, 79, 84–5, 97, 101–2, 109, 117
Weiner processes, xvi–xvii, xx–xxi, 37–44, 51–2, 67, 80–3, 85–6
  see also Brownian processes
Whole-yield curve models, concepts, 55, 56–7, 79, 97
Williams, D., 39
Wiseman, J., 104, 117

Yield, 10, 19, 55–6, 73, 80, 98, 147, 153–7
  Brady bonds, 169–74
  butterfly/barbell trades, 27–33, 28
  convexity bias, 155, 161–2
  coupon relationship, 10, 19
  default yield spread, 165–8
  determinants, 6–8, 19–23, 91

index-linked bonds, 141–2
interest rates, 57–61, 63–4
long bonds, 153–62
portfolios, 29–30
risks, 19–20
theoretical default spread, 165–8
UK benchmark-uses, 98
zero-coupon bonds, 73, 80, 98, 147, 153–60
Yield curves, xiii–xv, 6–11, 23–4, 79–92, 97–116, 117–39, 146–7
asset prices, xiii–xiv, 35–54
butterfly/barbell trades, 27–33
concepts, xiii–xv, 6–11, 23–4, 117–39, 146–7
convexity bias, 155, 161–2
coupon, 7–8, 10, 19, 24–7
cubic splines, 97, 101–2, 104–15, 117–39, 147–8
dynamics, xiv–xv, 23–4
estimation approaches, xiii, 97–116
fitting processes, 97–115, 117–39, 146–7
function estimates, 101–2
gilt market, 1, 5, 6–8
government authorities, 7–8, 97–9, 112–13
Heath–Jarrow–Morton model, xiv, 55, 56–7, 79–92, 97–8
historical analysis, 12–13
humped characteristics, 114–15, 137
influencing factors, xv, 23–4
information content, 98–102
inverted yield curves, 1, 7–8, 19–20, 23, 69
liquidity issues, 7–8, 13
local part of the yield curve, 7–8, 9–10, 23–4, 112–13
par yield curves, 8–11, 98
parametric fitting-techniques, 102–3, 108, 109–10, 114–15, 117–18, 136
polynomial approximations, 101–2, 104–5, 118–38
'ragged' curves, 101
relative value trading, 6–11, 19–33, 92–3
shapes, 1, 7–8, 19–20, 23, 69, 84–5, 114–15
short-term rates, xvi–xvii, 3–4
spline fitting-techniques, 97, 101–15, 117–39, 147–8
whole-yield curve models, 55, 56–7, 79, 97
Yield spread trades, 23–5, 163–5
background, 23–4
concepts, 23–4
www.YieldCurve.com, 132

Zero-coupon bonds, 55–6, 60–1, 69–75, 80–1, 84–7, 92, 97–115, 147, 154–7
Brady bonds, 169–71
long bonds, 153–4, 159–60
prices, 69–73, 75, 84–7, 154–9
yield, 73, 80, 98, 147, 153–60
Zero-coupon curve, xiv, xvii–xviii, 6–7, 8–11, 97–115, 119, 147, 155–7, 160–1
Zervos, D., 99, 104–9, 118, 136
Zheng, C., 97