Capital Market Instruments

Analysis and Valuation

Third Edition

Moorad Choudhry
Didier Joannas
Gino Landuyt
Richard Pereira
Rod Pienaar

Foreword by Vince Cable MP

Capital Market Instruments

Also by Moorad Choudhry

THE BOND AND MONEY MARKETS: Strategy, Trading, Analysis

STRUCTURED CREDIT PRODUCTS: Credit Derivatives and Synthetic Securitisation

THE GILT-EDGED MARKET

THE REPO HANDBOOK

ANALYSING AND INTERPRETING THE YIELD CURVE

THE FUTURES BOND BASIS

ADVANCED FIXED INCOME ANALYSIS

THE HANDBOOK OF EUROPEAN FIXED INCOME SECURITIES

THE HANDBOOK OF EUROPEAN STRUCTURED FINANCIAL PRODUCTS

BANK ASSET AND LIABILITY MANAGEMENT

THE CREDIT DEFAULT SWAP BASIS

Capital Market Instruments

Analysis and Valuation

3rd Edition

Moorad Choudhry

Didier Joannas

Gino Landuyt

Richard Pereira

Rod Pienaar

palgrave
macmillan

© Moorad Choudhry, Didier Joannas, Gino Landuyt, Richard Pereira and Rod Pienaar 2001, 2004, 2010
Foreword © Vincent Cable 2010

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No portion of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

Any person who does any unauthorized act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

The authors have asserted their rights to be identified as the authors of this work in accordance with the Copyright, Designs and Patents Act 1988.

First published by The Financial Times/Prentice Hall in 2001

Second edition published by Palgrave Macmillan in 2004

Published in paperback 2010 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC, 175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries

ISBN 978–0–230–57603–2 paperback

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

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

A catalog record for this book is available from the Library of Congress.

10 9 8 7 6 5 4 3 2 1
19 18 17 16 15 14 13 12 11 10

Printed and bound in Great Britain by
CPI Antony Rowe, Chippenham and Eastbourne

The views, thoughts and opinions expressed in this book are those of the author in his individual private capacity and should not in any way be attributed to Europe Arab Bank plc or Arab Bank Group, or to Moorad Choudhry as a representative, officer, or employee of Europe Arab Bank plc or Arab Bank Group.

The views, thoughts and opinions expressed in this book are those of the author in his individual capacity and should not in any way be attributed to Thomson Reuters or to Didier Joannas as a representative, officer, or employee of Thomson Reuters.

The views, thoughts and opinions expressed in this book are those of the author in his individual capacity and should not in any way be attributed to Europe Arab Bank plc or Arab Bank Group, or to Gino Landuyt as a representative, officer, or employee of Europe Arab Bank plc or Arab Bank Group.

Neither UBS AG nor any subsidiary or affiliate of UBS AG is in any way connected with the contents of this publication, which represents the independent work, conclusions and opinions of its authors. Accordingly while one of its authors, Rod Pienaar, is a current employee of UBS AG, no responsibility for loss occasioned to any person acting or refraining from action as a result of any statement in this publication can be accepted by either UBS AG or any subsidiary or affiliate of the UBS Group.

Dedicated to the beauty and timeless elegance of Lloyd Cole and the Commotions …

Moorad Choudhry

For Chloë and Max

Didier Joannas

'No steam or gas drives anything until it is confined. No life ever grows great until it is focused, dedicated and disciplined' (Harry Fosdick). Therefore I would like to dedicate this book to my mother as she has been the driving force behind these values which brought me to where I am today.

Gino Landuyt

For my parents, if I could choose any two then I'd stick with the two that I have.

Rod Pienaar

This page intentionally left blank

About the Authors

Moorad Choudhry is Head of Treasury at Europe Arab Bank plc in London. He was previously Head of Treasury at KBC Financial Products, and prior to that he was in structured finance with JPMorgan Chase, in bond proprietary trading at Hambros Bank and in gilt-edged market making and money market trading at ABN Amro Hoare Govett Sterling Bonds Limited. He began his City career at the London Stock Exchange in 1989.

Moorad is Visiting Professor at the Department of Economics, London Metropolitan University and Visiting Research Fellow at the ICMA Centre, University of Reading. He is a Fellow of the Securities and Investment Institute and is on the Editorial Board of the *Journal of Structured Finance*.

Didier Joannas is Regional Director for Thomson Reuters-Risk in North Asia. He was previously with SunGard Trading and Risk Systems in Hong Kong. He worked with Moorad Choudhry on the gilt-edged market making desk at ABN Amro Hoare Govett Sterling Bonds Limited and ABN Amro Securities (UK) Limited during the 1990s.

Didier obtained his doctorate in aerodynamics and aeronautical engineering at the University of St Etienne in France before joining Avions Marcel Dassault. The projects he was involved with at Dassault included the Ariane space rocket and the Rafaele supersonic jet fighter. He then worked for Viel & Cie futures brokers in Paris before joining ABN Amro Hoare Govett in London in 1994.

Gino Landuyt is Head of Treasury Sales at Europe Arab Bank plc in London. Prior to that he was Head of Mid-Cap Institutional Structured Sales at ING Bank in Brussels, and in sales and origination at KBC Bank NV in their Brussels, Frankfurt, New York and London offices.

Gino has an MA in Applied Economic Sciences from the State University of Antwerp, and an MBA from St. Ignatius University in Antwerp.

Francis Richard Pereira is an investment actuary specialising in Fixed Income and Alternative Investments. Richard worked for JPMorgan Chase as an Executive Director in the Structured Alternative Investments team, and prior to that he was a senior executive within the Fixed Income divisions of Nomura and Dresdner Kleinwort Wasserstein. Richard qualified as a Chartered Accountant at KPMG and is a Fellow of the Institute of Actuaries. He graduated from Imperial College, London with a First in Mathematics.

Rod Pienaar has 15 years of investment banking experience and is currently an Executive Director at UBS AG prime services in London. Prior to that he worked at Deutsche Bank AG prime services and in risk management consulting.

Contents

List of Figures xvii

List of Tables xxi

Foreword xxiii

Preface xxiv

Preface to the First Edition xxvi

Acknowledgements xxvii

PART I Introduction **1**

1 Introduction to Financial Market Instruments 3
Capital market financing 3
Derivative instruments 7
Securities and derivatives 9

2 Market-Determined Interest Rates, and the Time Value of Money 11
The market-determined interest rate 11
The time value of money 14

PART II Debt Capital Market Cash Instruments **23**

3 Money Market Instruments and Foreign Exchange 25
Overview 25
Securities quoted on a yield basis 26
Securities quoted on a discount basis 29
Commercial paper 31
Asset-backed commercial paper 35
Foreign exchange 39

4 Fixed Income Securities I: The Bond Markets 46
Introduction 46
Market participants 47
Bonds by issuers 49
The markets 56
Credit risk 60
Pricing and yield 63
Bond pricing and yield: the traditional approach 64
Accrued interest, clean and dirty bond prices 74
Bond pricing and yield: the current approach 77
Bond price in continuous time 81
Forward rates 86

	The term structure of interest rates	88
	Case study: Deriving a discount function	100
	Analysing and interpreting the yield curve	107
5	Fixed Income Securities II: Interest-Rate Risk	123
	Duration, modified duration and convexity	123
	Appendix 5.1: Measuring convexity	135
	Appendix 5.2: Taylor expansion of the price/yield function	137
6	Fixed Income Securities III: Credit Analysis, Bond Spreads and Relative Value Measurement	139
	Credit ratings	140
	Credit analysis	143
	Industry-specific analysis	147
	The art of credit analysis	149
	Bond spreads and relative value	150
	The Z-spread	152
7	Interest Rate Modelling	159
	Introduction	159
	One-factor term structure models	163
	Further one-factor term structure models	165
	The Heath, Jarrow and Morton model	167
	Choosing a term structure model	171
	Appendix 7.1: Geometric Brownian motion	173
8	Fitting the Yield Curve	176
	Yield curve smoothing	176
	Non-parametric methods	180
	Comparing curves	184
9	B-Spline Modelling and Fitting the Term Structure	186
	Introduction	186
	Bootstrapping	187
	An advanced methodology: the cubic B-spline	188
	Mathematical tools	193
	B-Splines	199
	Conclusion	201
10	Inflation-Indexed Bonds and Derivatives	203
	Introduction and basic concepts	203
	Index-linked bond yields	206
	Analysis of real interest rates	211
	Inflation-indexed derivatives	213
	Appendix 10.1: Current issuers of public-sector indexed securities	221
	Appendix 10.2: US Treasury inflation-indexed securities (TIPS)	221
11	Term Structure Construction Using QuantLib	224
	Getting started	224
	Using QuantLib's yield curve bootstrapping	225

PART III Structured Financial Products **235**

12 An Introduction to Asset-Backed Securities and Securitisation 237
The concept of securitisation 237
Reasons for undertaking securitisation 238
Benefits of securitisation to investors 240
The process of securitisation 240
Illustrating the process of securitisation 245
Securitisation post-credit crunch 249

13 Mortgage-Backed Securities 254
Mortgages 255
Mortgage risk 259
Securities 260
Cash flow patterns 261
Evaluation and analysis of mortgage-backed bonds 266

14 Collateralised Debt Obligations 273
An overview of CDOs 273
Investor analysis 277
Cash flow CDO structures 278
Synthetic CDOs 283
Risk and return on CDOs 287

PART IV Derivative Instruments **293**

15 Short-Term Interest-Rate Derivatives 295
Forward contracts 301
Short-term interest rate futures 302
Appendix 15.1: The forward interest rate and futures-implied forward rate 310
Appendix 15.2: Arbitrage proof of the futures price being equal to the forward price 311

16 Swaps 313
Interest rate swaps 313
Zero-coupon swap pricing 318
Non-vanilla interest-rate swaps 326
Currency swaps 328
Swaptions 331
An overview of interest-rate swap applications 335

17 Options I 341
Introduction 341
Option instruments 346
Option pricing: setting the scene 348

18 Options II 351
Option pricing 351
The Black–Scholes option model 353
Interest-rate options and the Black model 360
Comment on the Black–Scholes model 363
A final word on option models 364
Appendix 18.1: Summary of basic statistical concepts 366
Appendix 18.2: Lognormal distribution of returns 366
Appendix 18.3: The Black–Scholes model in Microsoft® Excel 367

19 Options III 370
Behaviour of option prices 370
Measuring option risk: the Greeks 372
The option smile 378
Caps and floors 381

20 Credit Derivatives 383
Introduction 383
Credit default swaps 388
Credit-linked notes 391
Total return swaps 392
CDS portfolio indices 398
Index tranche market 399
Impact of the 2007–8 credit crunch: new CDS contracts 402
General applications of credit derivatives 403
Credit derivatives pricing and valuation 409
Appendix 20.1: Sample term sheet for a credit default swap traded by XYZ Bank plc 423

PART V Equity Capital Markets 427

21 Introduction to Equity Instrument Analysis 429
Firm financial structure and company accounts 429
Valuation of shares 434
Dividend policy 437

22 Introduction to Financial Ratio Analysis 440
Key concepts in finance 440
Ratio analysis 443
Management-level ratio analysis 446
Corporate valuation 450
Appendix 22.1: Capital asset pricing model 451

PART VI Risk Measurement and Value-at-Risk 455

23 Value-at-Risk and Credit VaR 457
Introducing value-at-risk 457

Variance–covariance VaR 461
Historical VaR methodology 466
Simulation methodology 467
VaR for fixed income instruments 469
Stress testing 475
VAR methodology for credit risk 477
Modelling credit risk 477
CreditMetrics™ 479
Applications of credit VaR 484
Integrating the credit risk and market risk functions 486
Appendix 23.1: Assumption of normality 487

PART VII The 2007–9 Financial Market Crisis 489

24 Origins and Impact of a Financial Crisis 491
Chronology of a crisis: overview of the great credit crisis 2007–8 492
The shadow banking system 500
Globalisation, sovereign wealth funds and the Greenspan conundrum 502
The role of central banks 506
The carry trade and rise of derivatives 511
The rise in derivatives use and the role of 'quantitative finance' 514
Financial innovation and securitisation 517
Political interference 519
Further regulatory shortcomings 523
The credit rating agencies 526
Bad management practices 527
Conclusion and lessons to be learnt 528

Index 533

List of Figures

1.1 Financing supply and demand curves 5
2.1 The yield curve 13
2.2 Hypothetical discount function 21
3.1 Single-seller ABCP conduit 38
3.2 Multi-seller Euro ABCP conduit 38
4.1 Bloomberg screen IYC showing yield curves for US, UK, French and German government bond markets, 28 May 2009 52
4.2 Yield spread by rating and maturity 62
4.3 US Treasury zero-coupon yield curve, September 2000 79
4.4 UK gilt zero-coupon yield curve, September 2000 79
4.5 French OAT zero-coupon yield curve, September 2000 80
4.6 Par, spot and forward yield curves 97
4.7 Linear interpolation of money and future rates 101
4.8 Discount equation 103
4.9 Comparison of money market curves 104
4.10 The basic shapes of yield curves 108
4.11 Yield curve explained by expectations hypothesis and liquidity preference 114
5.1 lllustration of price sensitivity for three types of bonds, 15 December 2000 131
6.1 Bloomberg screen RATD 142
6.2 Bloomberg page IRSB for Pounds sterling, showing GBP swap rates and swap spread over UK gilts © Bloomberg L.P. All rights reserved. Reprinted with permission 151
6.3 Bloomberg page ASW for GKN bond, 10 August 2005 © Bloomberg L.P. All rights reserved. Reprinted with permission 153
6.4 Bloomberg page YAS for GKN bond, 10 August 2005 © Bloomberg L.P. All rights reserved. Reprinted with permission 154
6.5 Bloomberg page YAS for GKN bond, 10 August 2005 showing Z-spread history © Bloomberg L.P. All rights reserved. Reprinted with permission 155
6.6 Calculating the Z-spread, hypothetical 5% 2008 bond issued by XYZ plc 156
6.7 Bloomberg graph using screen G <go>, plot of asset-swap spread and CDS price for GKN bond, April–September 2005 © Bloomberg L.P. All rights reserved. Reprinted with permission 157
6.8 GKN bond, CDS basis during August–September 2005 158
7.1 Evolution of Brownian or Weiner process 174
8.1 Linear interpolation of bond yields, 9 November 2000 178
8.2 Spot and forward rates implied from rates in Figure 8.1 179

8.3 Cubic spline with knot points at 0, 2, 5, 10 and 25 years 181
9.1 Calculated discount function, spot and forward curves 194
10.1 The indexation lag 204
10.2 Synthetic index-linked bond 215
10.3 Year-on-year inflation swap 216
10.4 Illustration of a TIPS swap 216
10.5 Breakeven inflation swap 217
11.1 Yield curve screen 225
11.2 Market data screen 226
11.3 Day counter setting 226
11.4 Output setting 227
11.5 Futures splicing 228
11.6 Bootstrapping 229
11.7 Overlapping coupon 230
11.8 Interpolation 231
12.1 The process of securitisation 241
12.2 Amortising structures 243
12.3 The waterfall process 245
12.4 The effect on the liability side of the balance sheet of a securitisation transaction 246
12.5 Illustrative deal structure 247
13.1 Mortgage pass-through security with 0% constant prepayment rate 263
13.2 100% PSA model 263
13.3 200% PSA model 264
13.4 Modified duration and effective duration for agency mortgage-backed bonds 271
14.1 The CDO family 279
14.2 Generic cash flow CDO 280
14.3 Interest cash flow waterfall for cash flow CDO 281
14.4 Synthetic CDO structure 285
14.5 Total issuance by CDO underlying 288
14.6 Issuance data volumes, arbitrage CDO and balance sheet CDO 289
15.1 Key dates in an FRA trade 298
15.2 Rates used in FRA pricing 300
16.1 Cash flows for typical interest rate swap 317
16.2 Fixed-fixed rate currency swap 329
16.3 Bond issue with currency swap structure 330
16.4 Changing liability from floating to fixed-rate 336
16.5 Liability-linked swap, fixed to floating to fixed-rate exposure 336
16.6 Transforming a floating-rate asset to fixed-rate 337
16.7 PVBP of a 5-year swap and fixed-rate bond maturity period 339
17.1 Payoff profile for a bond futures contract 342
17.2 Payoff profile for call option contract 343
17.3 Basic option payoff profiles 345
19.1 Option lambda, nine-month bond option 379

19.2 (a) Bond option volatility smile 379
19.2 (b) Equity option volatility smile 380
20.1 Credit default swap notional volume outstanding 384
20.2 Credit default swap 388
20.3 Sovereign CDS levels during 2008–9 © Bloomberg L.P. Used with permission. Visit www.bloomberg.com 391
20.4 A cash-settled credit-linked note 392
20.5 Total return swap 393
20.6 Overview of CDS Portfolio Indices, by geographical region 398
20.7 Overview of Standardised Index Tranche Market 400
20.8 Illustration of tranched iTraxx Product 401
20.9 Bloomberg screen CDSW showing five-year CDS written on Virgin Media Finance plc name, traded 28 May 2009, with 'points upfront' valuation of 2.88% © Bloomberg L.P. Used with permission. Visit www.bloomberg.com 403
20.10 Bloomberg screen UPFR showing reference name recovery rates © Bloomberg L.P. Used with permission. Visit www.bloomberg.com 404
20.11 Reducing credit exposure 405
20.12 Haarman & Reimer loan description © Bloomberg L.P. Used with permission. 407
20.13 Chemicals sector CDS prices from Banco Bilbao Vizcaya, 9 March 2004 © Bloomberg L.P. © BBVA. Used with permission. 408
20.14 Probability of survival 421
20.15 Credit curves 422
22.1 Components of the cost of capital 452
23.1 The normal approximation of returns 469
23.2 Term structure used in the valuation 470
23.3 Comparison of distribution of market returns and credit returns 478
23.4 A binomial model of credit risk 480
23.5 Distribution of credit returns by rating 480
23.6 Analytics road map for CreditMetrics 481
23.7 Constructing the distribution value for a BBB-rated bond 483
23.8 Size of total exposure to obligor-risk/return profile 485
23.9 Leptokurtosis 488
24.1 Rise of sub-prime mortgage bond risk: the ABX index © Markit. Reproduced with permission 492
24.2 US Fed Fund rates: 2000–8 © Bloomberg L.P. Used with permission. Visit www.bloomberg.com 493
24.3 Sub-prime mortgage delinquency rates. 494
24.4 Asset-Backed Commercial Paper (ABCP) and non-ABCP market 495
24.5 TED Spread history, 2006–2009 © Bloomberg L.P. Used with permission. Visit www.bloomberg.com 496
24.6 Chronology of major events in the Credit Crisis 2007–8 498
24.7 Behaviour of iTraxx index price during main events © Bloomberg L.P. Used with permission. Visit www.bloomberg.com 499

24.8 Behaviour of S&P during main events © Bloomberg L.P. Used with permission. Visit www.bloomberg.com 499
24.9 A network of interest-rate swap arrangements 500
24.10 Relative share of total financial intermediary assets (in %) US data 501
24.11 Development in US current-accounts balance as a percentage of GDP 503
24.12 Assets under management in $ tln (2006) 504
24.13 Top ten asset managers in $ tln (2006) 504
24.14 US 10-year Treasury Note 2000–9 yield © Bloomberg L.P. Used with permission. Visit www.bloomberg.com 505
24.15 Net capital outflows from countries with current-account surpluses in $ bln 506
24.16 US two-year nominal GDP growth versus Fed fund rates 1960–2008 509
24.17 The US credit–output ratio 510
24.18 Growth of derivatives, 1998–2008 512
24.19 USD 3-month Libor versus USD 10-year swap rates 1999–2009 © Bloomberg L.P. Used with permission. Visit www.bloomberg.com 512
24.20 Credit Suisse/Tremont Hedge Fund Index Performance 514
24.21 The global OTC derivatives market 515
24.22 Concentration risk of OTC derivatives among US banks 516
24.23 Percentages of US home ownership, 1965–2007 521
24.24 Fed mortgage pools versus private RMBS 524
24.25 Weighting of causes of the Great Credit Crisis 2007–8 529

List of Tables

2.1	Hypothetical set of bonds and bond prices	20
2.2	Discount factors calculated using bootstrapping technique	20
3.1	Comparison of US CP and Eurocommercial CP	33
4.1	Government bond markets: characteristics of selected countries	50
4.2	Country yield curves (as of 28 May 2009)	51
4.3	Selected euro-denominated Eurobond issues in 2009	61
4.4	Hypothetical UK government bond yields as at 7 December 2000	90
4.5	Implied spot and forward rates	96
4.6	Money market rates	100
4.7	Discount factors	101
4.8	Forward discount factors	101
4.9	Zero-coupon discount factors	102
4.10	Discount factors	103
4.11	Positive yield curve with constant expected future rates (%)	117
5.1	Nature of the modified duration approximation	130
6.1	Summary of credit rating agency bond ratings	142
9.1	t_i values	192
9.2	UK gilt observed market and theoretical prices	193
15.1	Description of LIFFE short sterling future contract	303
16.1	Three-year cash flows	315
16.2	Swap quotes	318
16.3	Generic interest-rate swap	324
16.4	Generic interest-rate swap (Excel formulae)	324
16.5	Interest rate data for swaption valuation	334
16.6	PVBP for interest-rate swap	338
17.1	Basic option terminology	346
18.1	Microsoft Excel calculation of vanilla option price	367
19.1	Delta-neutral hedging for changes in underlying price	374
19.2	Dynamic hedging as a result of changes in volatility	377
20.1	Comparison of model results, expiry in six months	414
20.2	Comparison of model results, expiry in 12 months	414
21.1	Hypothetical corporate balance sheet	432
21.2	Hypothetical corporate profit and loss statement	433
22.1	Constructa plc balance sheet for the year ended 31 December 2000	444
22.2	Constructa plc profit and loss account for the year ended 31 December 2000	445
22.3	Constructa plc: notes to the accounts	445
22.4	Constructa plc RONA ratio measures	446
22.5	A UK plc corporate performance 1995–9	447
22.6	Constructa plc corporate-level ratios	447

22.7 Hypothetical company results 448
22.8 Gearing ratios 449
22.9 Comparable company financial indicators, year 2000 450
22.10 Peer group company ratios, mean values and Constructa plc market valuation 451
23.1 Two-asset portfolio VaR 461
23.2 Matrix variance–covariance calculation for two-asset portfolio 462
23.3 RiskMetrics grid points 464
23.4 Bond position to be mapped to grid points 465
23.5 Cash flow mapping and portfolio variance 465
23.6 Monte Carlo simulation results 468
23.7 Sample three-bond portfolio 470
23.8 Bond portfolio valuation 471
23.9 Bond portfolio undiversified VaR 472
23.10 Undiversified VaR for 3 v 6 FRA 472
23.11 Fixed-rate leg of five-year interest rate swap and undiversified VaR 474

Foreword

It gives me great pleasure to introduce this latest edition of Moorad Choudhry's textbook *Capital Market Instruments: Analysis and Valuation*.

The global financial crisis has exposed to wide public attention the extraordinary complexity of the financial instruments in which the financial services industry is now dealing. City traders – the Masters of the Universe – have stolen a march on the regulators whose job it was to scrutinise them. For that we have paid a significant price.

Instruments like SIVs, conduits, mortgage-backed securities and credit default swaps were treated as innovative, risk spreading, niche ideas, the preserve of pointy headed investment bankers, and not as the high-risk, potentially destructive gamble they have proved to be. The failure to grasp the importance of what was going on in the City or anticipate the potential flaws in these new mechanisms allowed a significant systemic risk to build.

It is imperative, if we are to prevent a repeat of the recent crisis, that the FSA has a fuller understanding of the kinds of deals that are being made in the City. There has been much debate, and will be much more, as to the proper way to manage the risk created by these complex instruments. Serious consideration should be given to splitting up banks so that everyday retail functions are protected from the risk of casino banking and derivatives must be moved away from over-the-counter trading into a central clearing house or onto an exchange that can be more easily monitored. Much is made of the importance of attracting the best and brightest talent to London (at the expense, it is suggested, of Frankfurt, Geneva and Hong Kong). But little is made of the counter-argument that such high pay for bankers deters talented individuals from pursuing careers in engineering, science and – most importantly in this context – financial regulation. I hope readers will recognise how considered regulation is a valuable and worthy contribution to a sound economy.

Gillian Tett, in her book *Fool's Gold*, has produced a thorough, journalistic account of the causes of the global financial crisis. *Capital Market Instruments* strength is as an analytical study of the instruments themselves, and while it is unlikely to appeal to the general reader, its aim – to provide a thorough understanding of financial markets – deserves a wide audience. It is not only bankers, regulators and politicians who need to understand the world of complex financial instruments. Journalists and academics also have a duty to comprehend and explain. This book provides an education as to where we are now, but the City has proved it can adapt to meet new challenges and it is vital, if further crisis is to be averted, that we continue to keep pace.

Vincent Cable
MP for Twickenham & Liberal Democrat Deputy Leader and
Shadow Chancellor of the Exchequer
7 October 2009

Preface

This book began life as a concise but complete textbook on financial market instruments and analysis, aimed at practitioners and graduate students. The third edition stays on this same course, but also includes a detailed assessment and analysis of the 2007–8 financial crisis or 'credit crunch'. Events during this period challenged banks and investors to re-evaluate erstwhile assumptions about the financial markets, which is why it is worth considering them in detail here. We have also updated chapters to account for the changes and developments that have occurred since the second edition was published. This is most significant in the chapter on credit derivatives, where among other things we describe the 'big bang' that occurred in April 2009, and which has resulted in North American contracts being traded under new cash flow arrangements.

Features of the third edition

This book is organised into seven parts. Part I sets the scene with a discussion on the financial markets, the time value of money and the determinants of the discount rate. Part II describes fixed income instruments, and the analysis and valuation of bonds. This covers in overview fashion the main interest-rate models, before looking in detail at some important areas of the markets, including

- fitting the yield curve, and an introduction to spline techniques
- the B-spline method of extracting the discount function
- bond pricing in continuous time
- inflation-indexed bonds.

We have removed the chapter on option-adjusted spread, as this is a specialist technique and better suited to a dedicated book on fixed income. There is a new chapter on bond credit analysis, and the various relative value measures used to assess bond return. Part II also has a new chapter on using QuantLib to construct a term structure model. This replaces the chapter describing the RATE yield curve application that was included in the previous edition.

Part III is an introduction to securitisation and structured financial products, with a look at mortgage-backed securities and collateralised debt obligations (CDOs). New material in this chapter includes a look at securitisation post-credit crunch, after banks started to undertake in-house deals in order to be able to raise funding at their central bank. The chapter on CDOs has been revised and updated.

In part IV we introduce the main analytical techniques used for derivative instruments. This includes futures and swaps, as well as an introduction to options and the Black-Scholes model, still widely used today over 30 years after its

introduction. Part V considers the basic concepts in equity analysis, using a hypothetical corporate entity for case study purposes. Part VI introduces the value-at-risk methodology, while the final part of the book is a new chapter assessing the causes of the 2007–8 financial market crisis.

Yield curve modelling application

In the first two editions of this book we included a specialist computer application, RATE, which was designed to introduce readers to yield curve modelling. For this edition we have decided to use www.quantlib.org (QuantLib), or more specifically its Excel add-in, for the purposes of demonstrating yield curve construction. QuantLib is a free library for quantitative finance. We do not provide a detailed description of what QuantLib does or how to use it because this can all be found on their web site, however Chapter 11 provides additional information on term structure construction, using QuantLib to demonstrate the concepts.

Preface to the First Edition

The book is a concise introduction to some of the important issues in financial market analysis, with an emphasis on fixed income instruments such as index-linked bonds, asset-backed securities, mortgage-backed securities, and related products such as credit derivatives. However, fundamental concepts in equity market analysis, foreign exchange and money markets, and certain other derivative instruments are also covered to complete the volume. The focus is on analysis and valuation techniques, presented for the purposes of practical application. Hence, institutional and market-specific data is largely omitted for reasons of space and clarity, as this is abundantly available in existing literature. Students and practitioners alike should be able to understand and apply the methods discussed here. The book attempts to set out a practical approach in presenting the main issues and the reader should benefit from the practical examples presented in the chapters. The material in the book has previously been used by the authors as a reference and guide on consulting projects at a number of investment banks worldwide.

The contents are aimed at those with a basic understanding of the capital markets; however this book also investigates the instruments to sufficient depth to be of use to the more experienced practitioner. It is primarily aimed at front office, middle office and back office staff working in banks and other financial institutions and who are involved to some extent in the capital markets. Undergraduate and postgraduate students of finance and economics should also find the presentation useful. Others including corporate and local authority treasurers, risk managers, capital market lawyers, auditors, financial journalists and professional students may find the broad coverage to be of value. In particular however, graduate trainees beginning their careers in financial services and investment banking should find the topic coverage ideal, as the authors have aimed to present the key concepts in both debt and equity capital markets.

Please note that to avoid needless repetition of 'he (or she)' in the text the term 'he' should be taken to indicate both the male and female gender.

Comments on the text are welcome and should be sent to the authors care of FT Prentice Hall.

Acknowledgements

Love, thanks and respect to Mum, Dad, Amileela, and Linzi. And to the *Raynes Park Footy Boys*.

Moorad Choudhry
Surrey, England
May 2009

Thanks to Moorad 'Goldfinger' Choudhry for making me part of this great adventure – for a third time!

Didier Joannas
Hong Kong
May 2009

Sitting around waiting for luck to come your way is as misguided as thinking that good things always come to those who 'want it enough'. The truth is that determination and desire are necessary but not sufficient. We have to try like crazy; we have to retain a relentless sense of determination; … and yet there are still no guarantees. Even after all that, we may come up empty-handed. That is the bleak but unavoidable logic of anyone who has deep ambitions.

—Ed Smith, *What Sport Tells Us About Life*, Penguin Viking 2008

PART I
Introduction

Part I of the book is a brief introduction to capital market instruments, designed to set the scene and discuss the concept of time value of money. There are a large number of text books that deal with the subjects of macroeconomics and corporate finance, and so these issues are not considered here. Instead we concentrate on the financial arithmetic that is the basic building block of capital market instruments analysis. We also consider briefly the determinants of interest rates or discount rates, which are key ingredients used in the valuation of capital market instruments.

This page intentionally left blank

1
Introduction to Financial Market Instruments

This book is concerned with the valuation and analysis of capital market securities, and associated derivative instruments, which are not *securities* as such but are often labelled thus. The range of instruments is large and diverse, and it would be possible to stock a library full of books on various aspects of this subject. Space dictates that the discussion be restricted to basic, fundamental concepts as applied in practice across commercial and investment banks and financial institutions around the world. The importance of adequate, practical and accessible methods of analysis cannot be overstated, as this assists greatly in maintaining an efficient and orderly financial system. By employing sound analytics, market participants are able to determine the fair pricing of securities, and thereby whether opportunities for profit or excess return exist.

In this chapter we define cash market securities and place them in the context of corporate financing and capital structure; we then define *derivative* instruments, specifically financial derivatives.

Capital market financing

In this section we briefly introduce the structure of the capital market, from the point of view of corporate financing. An entity can raise finance in a number of ways, and the flow of funds within an economy, and the factors that influence this flow, play an important part in the economic environment in which a firm operates. As in any market, pricing factors are driven by the laws of supply and demand, and price itself manifests itself in the *cost of capital* to a firm and the *return* expected by investors who supply that capital. Although we speak in terms of a corporate firm, many different entities raise finance in the capital markets. These include sovereign governments, supranational bodies such as the World Bank, local authorities and state governments, and public sector bodies or *parastatals*. However, equity capital funding tends to be the preserve of the firm.

Financing instruments

The key distinction in financing arrangements is between *equity* and *debt*. Equity finance represents ownership rights in the firm issuing equity, and may be raised

either by means of a share offer or as previous year profits invested as retained earnings. Equity finance is essentially permanent in nature, as it is rare for firms to repay equity; indeed in most countries there are legal restrictions to so doing.

Debt finance represents a loan of funds to the firm by a *creditor*. A useful way to categorise debt is in terms of its maturity. Hence very short-term debt is best represented by a bank overdraft or short-term loan, and for longer-term debt a firm can take out a bank loan or raise funds by issuing a bond. Bonds may be secured on the firm's assets or unsecured, or they may be issued against incoming cash flows, which is known as securitisation. The simplest type of bond is known as a *plain vanilla* or *conventional* bond, or in the US markets, a *bullet* bond. Such a bond features a fixed *coupon* and fixed term to maturity, so for example a US Treasury security such as the 3% 2015 pays interest on its nominal or face value of 3% each year until 15 August 2015, when it is redeemed and principal paid back to bondholders.

A firm's financing arrangements are specified in a number of ways, which include:

- The *term* or maturity: financing that is required for less than one year is regarded as short-term, and money market securities are short-term in this way. Borrowing between one year and 10 years is considered medium-term, while longer-dated requirements are regarded as long-term. There is permanent financing, for example preference shares.
- *Size* of funding: the amount of capital required.
- The *risk* borne by suppliers of finance and the *return* demanded by them as the cost of bearing such risk. The risk of all financial instruments issued by one issuer is governed by the state of the firm and the economic environment in which it operates, but specific instruments bear specific risks. Secured creditors are at less risk of loss compared to unsecured creditors, while the owners of equity (shareholders) are last in line for repayment of capital in the event of the winding-up of a company. The return achieved by the different forms of finance reflects the risk exposure each form represents.

A common observation[1] is that although shares and share valuation are viewed as very important in finance and finance text books, in actual cash terms they represent a minor source of corporate finance. Statistics[2] indicate that the major sources of funding are retained earnings and debt.

Market mechanism for determining financing price

In a free market economy, which apart from a handful of exceptions is now the norm for all countries around the world, the capital market exhibits the laws of supply and demand. This means that the market price of finance is brought into *equilibrium* by the price mechanism. A simple illustration of this is given in

[1] For example see Higson (1995) p. 181.
[2] Ibid., see the table on p. 180.

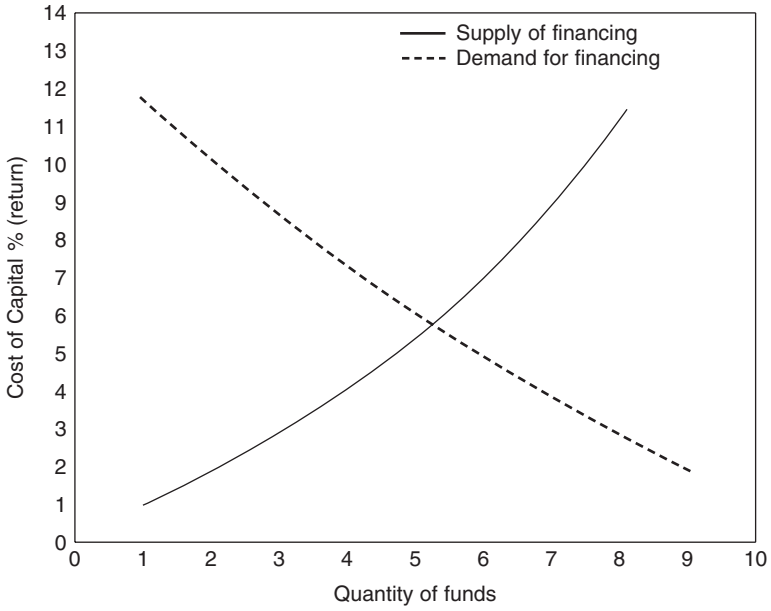


Figure 1.1 Financing supply and demand curves

Figure 1.1, which shows that the cost of finance will be the return level at which saving and investment are in equilibrium.

In Figure 1.1 the supply curve sloping upwards represents the investors' willingness to give up an element of present consumption when higher returns are available. The demand curve sloping downwards illustrates an increasing pipeline of projects that become more worthwhile as the cost of capital decreases. In the pioneering work of Fisher (1930) it was suggested that the cost of capital, in fact the rate of return required by the market, was made up of two components, the real return ri and the expected rate of inflation i. Extensive research since then has indicated that this is not the complete picture, for instance Fama (1975) showed that in the United States during the 1950s and 1960s, the change in the nominal level of interest rates was actually a reasonably accurate indicator of inflation, but that the real rate of interest remained fairly stable. Generally speaking the market's view on expected inflation is a major factor in driving nominal interest rates. On the other hand the real interest rate is generally believed[3] to be driven by factors that influence the total supply of savings and the demand for capital, which include overall levels of income and saving and government policies on issues such as personal and corporate taxation.

We look briefly at firm capital structure in Part V on equities.

[3] For example see Higson (1995), ch. 11.

Securities

The financial markets can be said to be an integration of market participants, the trading and regulatory environment (which includes stock and futures exchanges) and market instruments. These instruments can be further divided into cash securities and derivatives. Securities are known as cash market instruments (or simply *cash*) because they represent actual cash by value. A security product is issued by the party requiring finance, and as such represents a liability to the issuer. Conversely a security is an asset of the buyer or holder. Contrary to what might be thought given the publicity and literature emphasis on derivatives, financial markets are first and foremost cash securities, with the markets themselves being (in essence) a derivative of the wider economy.

In the first instance securities may be categorised as debt or equity. Such classification determines their ownership and participation rights with regard to the issuing entity. Generally speaking a holding of equity or *common stock* confers both ownership and voting rights. Debt securities do not confer such rights but rank ahead of equities in the event of a winding-up of the company.

Following this classification, securities are defined primarily in terms of their issuer, term to maturity (if not an equity) and currency. They may also be categorised in terms of:

- the rights they confer on the holder, such as voting and ownership rights
- whether they are unsecured or secured against fixed or floating assets
- the cash flows they represent
- how liquid they are, that is, the ease with which they can be bought and sold in the secondary market
- whether or not they offer a guaranteed return and/or redemption value
- the tax liability they represent
- their structure, for example if they are hybrid or composite securities, or whether their return or payoff profile is linked to another security.

The characteristics of any particular security influence the way it is valued and analysed. Debt securities originally were issued with an annual fixed interest or *coupon* liability, stated as a percentage of *par* value, so that their cash flows were known with certainty during their lifetime. This is the origin behind the term *fixed income* (or in sterling markets, *fixed interest*) security, although there are many different types of debt security issued that do not pay a fixed coupon. Equity does not pay a fixed coupon as the dividend payable is set each year, depending on the level of corporate after-tax profit for each year,[4] and even a dividend in time of profit is no longer obligatory. Witness the number of corporations that have never paid a dividend, such as Microsoft Corporation.[5]

[4] The exception is *preference shares* (in the United States, *preferred stock*), which combine certain characteristics of equity with others of debt.

[5] Given the performance of the company's share price since it was first listed, this fact is not likely to concern the owners of the shares too much.

Derivative instruments

In this book we consider the principal financial derivatives, which are forwards, futures, swaps and options. We also discuss briefly the importance of these instruments in the financial markets, and the contribution they have made to market efficiency and liquidity. Compared with a cash market security, a derivative is an instrument whose value is linked to that of an underlying asset. An example is a crude oil future, the value of which will track the value of crude oil. Hence the value of the future derives from that of the underlying crude oil. Financial derivatives are contracts written on financial securities or instruments, for example equities, bonds or other financial derivatives. In the following chapters we consider the main types of financial derivatives, namely forward contracts, futures, options and swaps. We do not deal with derivatives of other markets such as energy or weather, which are esoteric enough to warrant separate, specialist treatment.

Forward contracts

A forward contract is a tailor-made instrument, traded *over-the-counter* (OTC) directly between the counterparties, that is, agreed today for expiry at a point in the future. In the context of the financial markets a forward involves an exchange of an asset in return for cash or another asset. The price for the exchange is agreed at the time the contract is written, and is made good on delivery, irrespective of the value of the underlying asset at the time of contract expiry. Both parties to a forward are obliged to carry out the terms of the contract when it matures, which makes it different from an *option* contract.

Forward contracts have their origin in the agricultural commodity markets, and it is easy to see why this is so.[6] A farmer expecting to harvest his, say, wheat crop in four months' time is concerned that the price of wheat in four months might fall below the level it is at today. He can enter into a forward contract today for delivery when the crop is harvested; however the price the farmer receives will have been agreed today, so removing the uncertainty over what he will receive. The best known examples of forward contracts are forwards in foreign exchange (FX), which are in fact interest-rate instruments. A forward FX deal confirms the price today for a quantity of foreign currency that is delivered at some point in the future. The market in currency forwards is very large and liquid.

Futures contracts

Futures contracts, or simply futures, are exchange-traded instruments that are standardised contracts; this is the primary difference between futures and forwards. The first organised futures exchange was the Chicago Board of Trade, which opened for futures trading in 1861. The basic model of futures trading established in Chicago has been adopted around the world.

[6] See the footnote on page 10 of Kolb (2000), who also cites further references on the historical origin of financial derivatives.

Essentially futures contracts are standardised. That means each contract represents the same quantity and type of underlying asset. The terms under which delivery is made into an expired contract are also specified by the exchange. Traditionally futures were traded on an exchange's floor (in the 'pit') but this has been increasingly supplanted by electronic screen trading, so much so that by January 2004 the only trading floor still in use in London was that of the International Petroleum Exchange. The financial futures exchange, LIFFE, now trades exclusively on screen. Needless to say, the two exchanges in Chicago, the other being the Chicago Board Options Exchange, retained pit trading.

The differences between forwards and futures relate mainly to the mechanism by which the two instruments are traded. We have noted that futures are standardised contracts, rather than tailor-made ones. This means that they expire on set days of the year, and none other. Secondly, futures trade on an exchange, rather than OTC. Thirdly, the counterparty to every futures trade on the exchange is the exchange *clearing house*, which guarantees the other side to every transaction. This eliminates counterparty risk, and the clearing house is able to provide guarantees because it charges all participants a margin to cover their trade exposure. *Margin* is an initial deposit of cash or risk-free securities by a trading participant, plus a subsequent deposit to account for any trading losses, made at the close of each business day. This enables the clearing house to run a default fund. Although there are institutional differences between futures and forwards, the valuation of both instruments follows similar principles.

Swap contracts

Swap contracts are derivatives that exchange one set of cash flows for another. The most common swaps are interest-rate swaps, which exchange (for a period of time) fixed-rate payments for floating-rate payments, or floating-rate payments of one basis for floating payments of another basis.

Swaps are OTC contracts and so can be tailor-made to suit specific requirements. These requirements can be in regard to nominal amount, maturity or level of interest rate. The first swaps were traded in 1981 and the market is now well developed and liquid. Interest-rate swaps are so common as to be considered 'plain vanilla' products, similar to the way fixed-coupon bonds are viewed.

Option contracts

The fourth type of derivative instrument is fundamentally different from the other three products we have just introduced. This is because its payoff profile is unlike those of the other instruments, because of the optionality element inherent in the instrument. The history of options also goes back a long way. However, the practical use of financial options is generally thought of as dating from after the introduction of the acclaimed Black–Scholes pricing model for options, which was first presented by its authors in 1973.

The basic definition of option contracts is well known. A *call option* entitles its holder to buy the underlying asset at a price and time specified in the contract terms, the price specified being known as the *strike* or *exercise* price, while a

put option entitles its holder to sell the underlying asset. A European option can only be exercised on maturity, while an American option may be exercised by its holder at any time from the time it is purchased to its expiry. The party that has sold the option is known as the writer and its only income is the price or *premium* that it charges for the option. This premium should in theory compensate the writer for the risk exposure it is taking on when it sells the option. The buyer of the option has a risk exposure limited to the premium he paid. If a call option strike price is below that of the underlying asset price on expiry it is said to be *in-the-money*, otherwise it is *out-of-the-money*. When they are first written or struck, option strike prices are often set at the current underlying price, which is known as *at-the-money*.

For an excellent and accessible introduction to options we recommend Kolb (2000).

Securities and derivatives

Securities are commonly described as *cash* instruments because they represent actual cash, so that a 5% 10-year £100 million corporate bond pays 5% on the nominal value each year, and on maturity the actual nominal value of £100 million is paid out to bond holders. The risk to holders is potentially the entire nominal value or *principal* if the corporate entity defaults on the loan. Generally the physical flow of cash is essential to the transaction, for example when an entity wishes to raise finance. For other purposes, such as hedging or speculation, physical cash flow is not necessarily essential and the objectives can be achieved with non-cash or off-balance sheet instruments. The amount at risk in a derivative transaction is usually, but not always, considerably less than its nominal value. The use of derivatives can provide users with near-identical exposures to those in the cash market, such as changes in foreign exchange rates, interest rates or equity indices, but with reduced or no exposure to the principal or nominal value.

For instance a position in a 10-year £100 million sterling interest-rate swap has similar exposure to a position in the 10-year bond mentioned above, in terms of profit or loss arising from changes in sterling interest rates. However if the bond issuer is declared bankrupt, potentially the full value of the bond may be lost, whereas (if the same corporate is the swap counterparty) the loss for the swap holder would be considerably less than £100 million. As the risk with derivatives is lower than that for cash instruments (with the exception of writers of options), the amount of capital allocation required to be set aside by banks' trading derivatives is considerably less than that for cash. This is a key reason behind the popularity of derivatives, together with their flexibility and liquidity. The issue of banking capital is a particularly topical one, as the rules governing it were recently reformed; interested readers should consult Choudhry (2007).

In the next chapter we consider the basic building blocks of finance, the determination of interest rates and the time value of money.

Selected bibliography and references

Choudhry, M. *Bank Asset–Liability Management*, Wiley Asia, 2007.
Fama, E. 'Short-term interest rates as predictors of inflation', *American Economic Review*, 1975, pp. 269–82.
Fisher, I. *The Theory of Interest*, Macmillan, 1930.
Higson, C. *Business Finance*, 2nd edn, Butterworths, 1995.
Kolb, R. *Futures, Options and Swaps*, 3rd edn, Blackwell, 2000.
Van Horne, J. *Financial Management and Policy*, 10th edn, Prentice Hall, 1995.

2
Market-Determined Interest Rates, and the Time Value of Money

For any application the discount rate that is used is the market-determined rate. This rate is used to value capital market instruments. The rate of discount reflects the fact that cash has a current value and any decision to forgo consumption of cash today must be compensated at some point in the future. So when a cash-rich individual or entity decides to invest in another entity, whether by purchasing the latter's equity or debt, he is forgoing the benefits of consuming a known value of cash today for an unknown value at some point in the future. That is, he is sacrificing consumption today for the (hopefully) greater benefits of consumption later. The investor will require compensation for two things; first, for the period of time that his cash is invested and therefore unusable, and secondly for the risk that his cash may fall in value or be lost entirely during this time. The beneficiary of the investment, who has issued shares or bonds, must therefore compensate the investor for bearing these two risks. This makes sense, as if compensation was not forthcoming the investor would not be prepared to part with his cash.

The compensation payable to the investor is available in two ways. The first is through the receipt of cash income, in the form of interest income if the investment is in the form of a loan or a bond, dividends from equity, rent from property and so on, and the second is through an increase in the value of the original capital over time. The first is interest return or gain and the second is capital gain. The sum of these two is the overall rate of return on the investment.

The market-determined interest rate

The rate of interest

The interest rate demanded in return for an investment of cash can be considered the required rate of return. In an economist's world of no inflation and no default or other risk, the real interest rate demanded by an investor would be the equilibrium rate at which the supply of funds available from investors meets the demand for funds from entrepreneurs. The time preference of individuals determines whether they will be borrowers or lenders, that is, whether they wish to consume now or invest for consumption later. As this is not an economics textbook, we will

not present even an overview analysis; however the rate of interest at which both borrowing and lending takes place will reflect the time preference of individuals.

Assume that the interest rate is 4%. If this is too low, there will be a surplus of people who wish to borrow funds over those who are willing to lend. If the rate was 6% and this was considered too high, the opposite would happen, as there would be an excess of lenders over borrowers. The *equilibrium rate of interest* is that rate at which there is a balance between the supply of funds and the demand for funds.[1] The interest rate is the return received from holding cash or money, or the cost of *credit,* the price payable for borrowing funds. Sometimes the term *yield* is used to describe this return.

The rate of inflation

The equilibrium rate of interest would be the rate observed in the market in an environment of no inflation and no risk. In an inflationary environment, the compensation paid to investors must reflect the expected level of inflation. Otherwise, borrowers would be repaying a sum whose real value was being steadily eroded. We illustrate this in simple fashion.

Assume that the markets expect that the general level of prices will rise by 3% in one year. An investor forgoing consumption of £1 today will require a minimum of £1.03 at the end of a year, which is the same value (in terms of purchasing power) that he had at the start. His total rate of return required will clearly be higher than this, to compensate for the period of time when the cash was invested. Assume further then that the equilibrium real rate of interest is 2.50%. The total rate of return required on an investment of £1 for one year is calculated as:

$$\begin{aligned}\text{Repayment of principal} &= £1 \times (1 + \text{real interest rate}) \times \frac{\text{Price level at year-end}}{\text{Price level at start of year}} \\ &= £1 \times (1.025) \times \frac{£1.03}{£1} \\ &= £1\ (1.025)(1.03) \\ &= £1.05575\end{aligned}$$

or 5.575%. This is known as the *nominal* rate of interest. The nominal interest rate is determined using the Fisher equation after Fisher (1930) and is shown as (2.1).

$$\begin{aligned}1 + r &= (1 + \rho)(1 + i) \\ &= 1 + \rho + i + \rho i\end{aligned} \tag{2.1}$$

where

r is the nominal rate of interest
ρ is the real rate of interest
i is the expected rate of inflation and is given by

[1]There is of course not one interest rate, but many different interest rates. This reflects the different status of individual borrowers and lenders in a capital market.

$$i = \frac{\text{price level at end of period}}{\text{price level at start of period}} - 1$$

A market-determined interest rate must also account for what is known as the *liquidity premium*, which is the price paid for the conflict of interest between borrowers who wish to borrow (at preferably fixed rates) for as long a period as possible, and lenders who wish to lend for as short a period as possible. A short-dated instrument is generally more easy to transact in the secondary market than a long-dated instrument, that is, it is more liquid. The trade-off is that in order to entice lenders to invest for longer time periods, a higher interest rate must be offered. Combined with investors' expectations of inflation, this means that rates of return (or yields) are generally higher for longer-dated investments. This manifests itself most clearly in an upward sloping *yield curve*. Yield curves are considered in a later chapter; in Figure 2.1 we show a hypothetical upward sloping yield curve with the determinants of the nominal interest rate indicated.

Figure 2.1 shows two curves. The lower one incorporates the three elements we have discussed, those of the real rate, expected inflation and liquidity. However it would only apply for investments that bore no *default* risk, that is no risk that the borrower would default on the loan and not repay it.[2] Investments that are default-free are typified by government bonds issued by countries with developed economies, for example US Treasury securities or UK gilts. Investments that expose

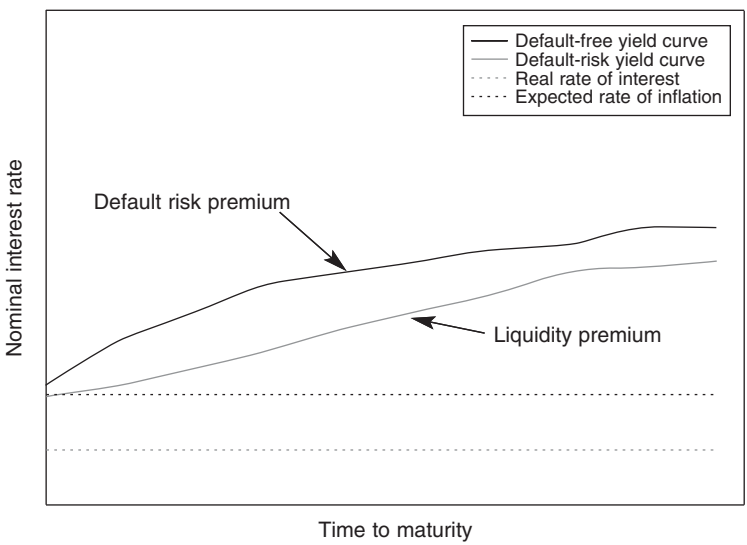


Figure 2.1 The yield curve

[2]The borrower may be unable to repay it, say because of bankruptcy or liquidation, or unwilling to repay it, for example due to war or revolution.

the investor to default risk, for example a corporate bond, must offer a return that incorporates a risk premium, over and above the risk-free interest rate. If this were not the case, investors would be reluctant to enter into such investments. The risk premium factor is indicated by the higher yield curve in Figure 2.1.

The time value of money

Present value and discounting

We now review a key concept in cash flow analysis, that of discounting and present value. It is essential to have a firm understanding of the main principles summarised here before moving on to other areas. When reviewing the concept of the time value of money, we assume that the interest rates used are the market determined rates of interest.

Financial arithmetic has long been used to illustrate that £1 received today is not the same as £1 received at a point in the future. Faced with a choice between receiving £1 today or £1 in one year's time we would not be indifferent given a rate of interest of say 10%, which was equal to our required nominal rate of interest. Our choice would be between £1 today or £1 plus 10p – the interest on £1 for one year at 10% per annum. The notion that money has a time value is a basic concept in the analysis of financial instruments. Money has time value because of the opportunity to invest it at a rate of interest. A loan that has one interest payment on maturity is accruing *simple interest*. On short-term instruments there is usually only the one interest payment on maturity, hence simple interest is received when the instrument expires. The terminal value of an investment with simple interest is given by:

$$F = P\,(1 + r) \tag{2.2}$$

where

F is the terminal value or *future value*
P is the initial investment or *present value*
r is the interest rate

The market convention is to quote interest rates as *annualised* interest rates, which is the interest that is earned if the investment term is one year. Consider a three-month deposit of £100 in a bank, placed at a rate of interest of 6%. In such an example the bank deposit will earn 6% interest for a period of 90 days. As the annual interest gain would be £6, the investor will expect to receive a proportion of this, which is calculated below:

$$£6.00 \times \frac{90}{365}$$

Therefore the investor will receive £1.479 interest at the end of the term. The total proceeds after the three months is therefore £100 plus £1.479. If we wish to

calculate the terminal value of a short-term investment that is accruing simple interest we use the following expression:

$$F = P\left(1 + r \times \frac{\text{days}}{\text{year}}\right) \tag{2.3}$$

The fraction $\frac{\text{days}}{\text{year}}$ refers to the numerator, which is the number of days the investment runs, divided by the denominator which is the number of days in the year. In the sterling markets the number of days in the year is taken to be 365, however most other markets (including the dollar and euro markets) have a 360-day year convention. For this reason we simply quote the expression as 'days' divided by 'year' to allow for either convention.

Let us now consider an investment of £100 made for three years, again at a rate of 6%, but this time fixed for three years. At the end of the first year the investor will be credited with interest of £6. Therefore for the second year the interest rate of 6% will be accruing on a principal sum of £106, which means that at the end of Year 2 the interest credited will be £6.36. This illustrates how *compounding* works, which is the principle of earning interest on interest. The outcome of the process of compounding is the *future value* of the initial amount. The expression is given in (2.4):

$$FV = PV\,(1 + r)^n \tag{2.4}$$

where

FV is the future value
PV is initial outlay or *present value*
r is the periodic rate of interest (expressed as a decimal)
n is the number of periods for which the sum is invested

When we compound interest we have to assume that the reinvestment of interest payments during the investment term is at the same rate as the first year's interest. That is why we stated that the 6% rate in our example was fixed for three years. We can see however that compounding increases our returns compared with investments that accrue only on a simple interest basis.

Now let us consider a deposit of £100 for one year, at a rate of 6% but with quarterly interest payments. Such a deposit would accrue interest of £6 in the normal way but £1.50 would be credited to the account every quarter, and this would then benefit from compounding. Again assuming that we can reinvest at the same rate of 6%, the total return at the end of the year will be:

$$100 \times [(1 + 0.015) \times (1 + 0.015) \times (1 + 0.015) \times (1 + 0.015)] = 100 \times [(1 + 0.015)^4]$$

which gives us 100×1.06136, a terminal value of £106.136. This is some 13 pence more than the terminal value using annual compounded interest.

In general if compounding takes place m times per year, then at the end of n years mn interest payments will have been made and the future value of the principal is given by (2.5).

$$FV = PV\left(1+\frac{r}{m}\right)^{mn} \tag{2.5}$$

As we showed in our example the effect of more frequent compounding is to increase the value of the total return compared with annual compounding. The effect of more frequent compounding is shown below, where we consider the annualised interest rate factors, for an annualised rate of 6%.

$$\text{Interest rate factor} = \left(1+\frac{r}{m}\right)^{m}$$

Compounding frequency	*Interest rate factor*	
Annual	$(1+r)$	$= 1.060000$
Semi-annual	$\left(1+\frac{r}{2}\right)^{2}$	$= 1.060900$
Quarterly	$\left(1+\frac{r}{4}\right)^{4}$	$= 1.061364$
Monthly	$\left(1+\frac{r}{12}\right)^{12}$	$= 1.061678$
Daily	$\left(1+\frac{r}{365}\right)^{365}$	$= 1.061831$

This shows us that the more frequent the compounding, the higher the interest rate factor. The last case also illustrates how a limit occurs when interest is compounded continuously. Equation (2.5) can be rewritten as follows:

$$\begin{aligned} FV &= PV\left[\left(1+\frac{r}{m}\right)^{m/r}\right]^{m} \\ &= PV\left[\left(1+\frac{1}{m/r}\right)^{m/r}\right]^{m} \\ &= PV\left[\left(1+\frac{1}{n}\right)^{n}\right]^{m} \end{aligned} \tag{2.6}$$

where $n = m/r$. As compounding becomes continuous and m and hence n approach infinity, the expression in the square brackets in (2.6) approaches a value known

as e, which is shown below.

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n} = 2.718281\ldots$$

If we substitute this into (2.6) this gives us:

$$FV = PVe^{rn} \tag{2.7}$$

where we have continuous compounding. In (2.7) e^{rn} is known as the *exponential function* of rn and it tells us the continuously compounded interest rate factor. If $r = 6\%$ and $n = 1$ year, then:

$$e^{r} = (2.718281)^{0.05} = 1.061837$$

This is the limit reached with continuous compounding.

The convention in both wholesale and personal (retail) markets is to quote an annual interest rate. A lender who wishes to earn the interest at the rate quoted has to place her funds on deposit for one year. Annual rates are quoted irrespective of the maturity of a deposit, from overnight to ten years or longer. For example, if one opens a bank account that pays interest at a rate of 3.5% but then closes it after six months, the actual interest earned will be equal to 1.75% of the sum deposited. The actual return on a three-year building society bond (fixed deposit) that pays 6.75% fixed for three years is 21.65% after three years. The quoted rate is the annual one-year equivalent. An overnight deposit in the wholesale or *interbank* market is still quoted as an annual rate, even though interest is earned for only one day.

The convention of quoting annualised rates is to allow deposits and loans of different maturities and different instruments to be compared on the basis of the interest rate applicable. We must be careful when comparing interest rates for products that have different payment frequencies. As we have seen from the foregoing paragraphs, the actual interest earned will be greater for a deposit earning 6% on a semi-annual basis than for one earning 6% on an annual basis. The convention in the money markets is to quote the equivalent interest rate applicable when taking into account an instrument's payment frequency.

We saw how a *future value* could be calculated given a known *present value* and rate of interest. For example £100 invested today for one year at an interest rate of 6% will generate $100 \times (1 + 0.06) = £106$ at the end of the year. The future value of £100 in this case is £106. We can also say that £100 is the *present value* of £106 in this case.

In equation (2.4) we established the following future value relationship:

$$FV = PV\,(1 + r)^{n}$$

By reversing this expression we arrive at the present value calculation given in (2.8).

$$PV = \frac{FV}{(1+r)^{n}} \tag{2.8}$$

where the symbols represent the same terms as before. Equation (2.8) applies in the case of annual interest payments, and enables us to calculate the present value of a known future sum.

To calculate the present value for a short-term investment of less than one year we will need to adjust what would have been the interest earned for a whole year by the proportion of days of the investment period. Rearranging the basic equation, we can say that the present value of a known future value is:

$$PV = \frac{FV}{\left(1 + r \times \frac{\text{days}}{\text{years}}\right)} \tag{2.9}$$

Given a present value and a future value at the end of an investment period, what then is the interest rate earned? We can rearrange the basic equation again to solve for the *yield*.

When interest is compounded more than once a year, the formula for calculating present value is modified, as shown in (2.10).

$$PV = \frac{FV}{\left(1 + \frac{r}{m}\right)^{mn}} \tag{2.10}$$

where as before FV is the cash flow at the end of year n, m is the number of times a year interest is compounded, and r is the rate of interest or discount rate. Illustrating this therefore, the present value of £100 that is received at the end of five years at a rate of interest rate of 5%, with quarterly compounding is:

$$\begin{aligned} PV &= \frac{100}{\left(1 + \frac{0.05}{4}\right)^{(4)(5)}} \\ &= £78.00 \end{aligned}$$

Interest rates in the money markets are always quoted for standard maturities, for example overnight, 'tom next' (the overnight interest rate starting tomorrow, or 'tomorrow to the next'), spot next (the overnight rate starting two days forward), one week, one month, two months and so on up to one year. If a bank or corporate customer wishes to deal for non-standard periods, an interbank desk will calculate the rate chargeable for such an 'odd date' by *interpolating* between two standard period interest rates. If we assume that the rate for all dates in between two periods increases at the same steady state, we can calculate the required rate using the formula for *straight line* interpolation, shown in (2.11).

$$r = r_1 + (r_2 - r_1) \times \frac{n - n_1}{n_2 - n_1} \tag{2.11}$$

where

r is the required odd-date rate for n days
r_1 is the quoted rate for n_1 days
r_2 is the quoted rate for n_2 days

Let us imagine that the one-month (30-day) offered interest rate is 5.25% and that the two-month (60-day) offered rate is 5.75%.[3] If a customer wishes to borrow money for a 40-day period, what rate should the bank charge? We can calculate the required 40-day rate using the straight line interpolation process. The increase in interest rates from 30 to 40 days is assumed to be 10/30 of the total increase in rates from 30 to 60 days. The 40-day offered rate would therefore be:

$$5.25\% + (5.75\% - 5.25\%) \times 10/30 = 5.4167\%$$

What about the case of an interest rate for a period that lies just before or just after two known rates and not roughly in between them? When this happens we extrapolate between the two known rates, again assuming a straight line relationship between the two rates and for a period after (or before) the two rates. So if the one-month offered rate is 5.25% while the two-month rate is 5.75%, the 64-day rate is:

$$5.25 + (5.75 - 5.25) \times 34/30 = 5.8167\%$$

Discount factors

An n-period discount factor is the present value of one unit of currency (£1 or $1) that is payable at the end of period n. Essentially it is the present value relationship expressed in terms of £1. If $d(n)$ is the n-year discount factor, then the five-year discount factor at a discount rate of 6% is given by:

$$d(5) = \frac{1}{(1+0.06)^5} = 0.747258$$

The set of discount factors for every time period from one day to 30 years or longer is termed the *discount function*. Discount factors may be used to price any financial instrument that is made up of a future cash flow. For example what would be the value of £103.50 receivable at the end of six months if the six-month discount factor is 0.98756? The answer is given by:

$$0.98756 \times 103.50 = 102.212$$

In addition discount factors may be used to calculate the future value of any present investment. From the example above, £0.98756 would be worth £1 in six months' time, so by the same principle a present sum of £1 would be worth

$$1/d(0.5) = 1/0.98756 = 1.0126$$

at the end of six months.

[3] This is the convention in the sterling and euro markets that is, 'one month' is 30 days.

Table 2.1 Hypothetical set of bonds and bond prices

Coupon	*Maturity date*	*Price*
7%	07 June 01	101.65
8%	07 December 01	101.89
6%	07 June 02	100.75
6.50%	07 December 02	100.37

It is possible to obtain discount factors from current bond prices. Assume a hypothetical set of bonds and bond prices as given in Table 2.1, and assume further that the first bond in the table matures in precisely six months' time (these are semi-annual coupon bonds).

Taking the first bond, this matures in precisely six months' time, and its final cash flow will be £103.50, comprised of the £3.50 final coupon payment and the £100 redemption payment. The price or present value of this bond is £101.65, which allows us to calculate the six-month discount factor as:

$$d(0.5) \times 103.50 = 101.65$$

which gives $d(0.5)$ equal to 0.98213.

From this first step we can calculate the discount factors for the following six-month periods. The second bond in Table 2.1, the 8% 2001, has the following cash flows:

£4 in six months' time
£104 in one year's time

The price of this bond is £101.89, which again is the bond's present value, and this consists of the sum of the present values of the bond's total cash flows. So we are able to set the following:

$$101.89 = 4 \times d(0.5) + 104 \times d(1)$$

However we already know $d(0.5)$ to be 0.98213, which leaves only one unknown in the above expression. Therefore we may solve for $d(1)$ and this is shown to be 0.94194.

If we carry on with this procedure for the remaining two bonds, using successive discount factors, we obtain the complete set of discount factors as shown in Table 2.2.

Table 2.2 Discount factors calculated using bootstrapping technique

Coupon	*Maturity date*	*Term (years)*	*Price*	*d(n)*
7%	07 June 01	0.5	101.65	0.98213
8%	07 December 01	1.0	101.89	0.94194
6%	07 June 02	1.5	100.75	0.92211
6.5%	07 December 02	2.0	100.37	0.88252

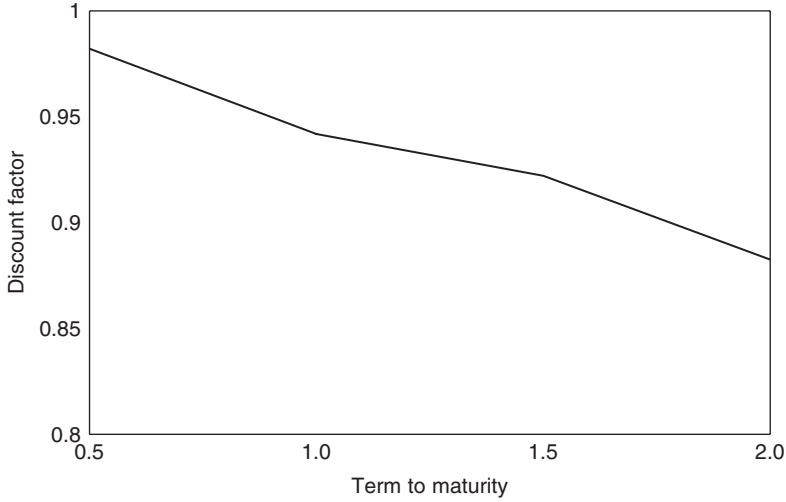


Figure 2.2 Hypothetical discount function

The continuous function for the two-year period from today is shown as the discount function, in Figure 2.2.

This technique, which is known as *bootstrapping*, is conceptually neat, but problems arise when we do not have a set of bonds that mature at precise six-month intervals. In addition liquidity issues connected with specific individual bonds can also cause complications. However it is still worth being familiar with this approach.

Note from Figure 2.2 how discount factors decrease with increasing maturity: this is intuitively obvious, since the present value of something to be received in the future diminishes the further into the future we go.

Selected bibliography and references

Blake, D. *Financial Market Analysis*, McGraw Hill, 1990.
Fisher, I. *Theory of Interest*, Macmillan, 1930.
Lee, T. *Economics for Professional Investors*, 2nd edn, Prentice Hall, 1998.

This page intentionally left blank

PART II
Debt Capital Market Cash Instruments

Part II of this book concentrates on vanilla debt market instruments. We begin with money market instruments. The first products in any capital market are money market instruments such as Treasury bills and bankers' acceptances. These, together with other cash money market products, are considered in Chapter 3. The next three chapters are devoted to fixed-income instruments or bonds. The analysis generally restricts itself to default-free bonds. Chapter 4 is a large one, which begins by describing bonds in the 'traditional' manner, and then follows with the current style of describing the analysis of bonds. There is also a description of the bootstrapping technique of calculating spot and forward rates. In Chapter 7 we summarise some of the most important interest-rate models used in the market today. This is a well-researched topic and the bibliography for this chapter is consequently quite sizeable. The crux of the analysis presented is the valuation of future cash flows. We consider the pricing of cash flows whose future value is known, in intermediate-level terms. The reader requires an elementary understanding of statistics, probability and calculus to make the most of these chapters. There are a large number of texts that deal with the mathematics involved; an overview of these is given in Choudhry (2005).

We look first at default-free zero-coupon bonds. The process begins with the fair valuation of a set of cash flows. If we are analysing a financial instrument comprised of a cash flow stream of nominal amount C_i, paid at times $i = 1, 2, \ldots, N$ then the value of this instrument is given by:

$$PV = \sum_{i=1}^{N} C_i P(0, t_i)$$

where $P(0, t_i)$ is the price today of a zero-coupon bond of nominal value 1 maturing at each point i, or in other words the i-period discount factor. This expression can be written as:

$$PV = \sum_{i=1}^{N} C_i \exp[-(t_i) r(0, t_i)]$$

which indicates that in a no-arbitrage environment the present value of the cash flow stream is obtained by discounting the set of cash flows and summing them. Therefore in theory it is straightforward to calculate the present value of any cash flow stream (and by implication virtually any financial instrument) using the yields observed on a set of risk-free and default-free zero-coupon bonds.

In a market where such default-free zero-coupon bonds existed for all maturities, it would be relatively straightforward to extract the discount function to the longest-dated maturity, and we could use this discount function to value other cash flows and instruments. However, this is a theoretical construct because in practice there is no market with such a preponderance of risk-free zero-coupon bonds; indeed zero-coupon bonds are a relative rarity in government markets around the world. In practice, the set of such zero-coupon bonds is limited and is influenced by liquidity and other market considerations. We require therefore an efficient and tractable method for extracting the zero-coupon yield curve from coupon-paying bonds of varying maturity. This vital issue is introduced in Chapter 8, and is followed in Chapter 9 by an advanced-level treatment of the B-spline method of extracting the discount function. This is a most efficient technique.

Chapter 10 in Part II considers the analysis of inflation-indexed bonds, an important asset class in a number of capital markets around the world.

Reference

Choudhry, M. *Fixed Income Markets*, Wiley Asia, 2005.

3
Money Market Instruments and Foreign Exchange

Money market securities are debt securities with maturities of up to 12 months. Market issuers include sovereign governments, which issue Treasury bills, corporates issuing commercial paper, and banks issuing bills and certificates of deposit. Investors are attracted to the market because the instruments are highly liquid and carry relatively low credit risk. Investors in the money market include banks, local authorities, corporations, money market investment funds and individuals. However the money market is essentially a wholesale market and the denominations of individual instruments are relatively large.

In this chapter we review the cash instruments traded in the money market as well as the two main money market derivatives, interest-rate futures and forward-rate agreements.

Overview

The cash instruments traded in the money market include the following:

- Treasury bill
- time deposit
- certificate of deposit
- commercial paper
- bankers acceptance
- bill of exchange.

We can also add the market in repurchase agreements or *repo*, which are essentially secured cash loans, to this list.

A Treasury bill is used by sovereign governments to raise short-term funds, while certificates of deposit (CDs) are used by banks to raise finance. The other instruments are used by corporates and occasionally banks. Each instrument represents an obligation on the borrower to repay the amount borrowed on the maturity date, together with interest if this applies. The instruments above fall into one of two main classes of money market securities: those quoted on a *yield* basis and those quoted on a *discount* basis. These two terms are discussed below.

The calculation of interest in the money markets often differs from the calculation of accrued interest in the corresponding bond market. Generally the day-count convention in the money market is the exact number of days that the instrument is held over the number of days in the year. In the sterling market the year base is 365 days, so the interest calculation for sterling money market instruments is given by (3.1).

$$i = \frac{n}{365} \tag{3.1}$$

The majority of currencies including the US dollar and the euro calculate interest based on a 360-day base.

Settlement of money market instruments can be for value today (generally only when traded in before midday), tomorrow or two days forward, known as *spot*.

Securities quoted on a yield basis

Two of the instruments in the list above are yield-based instruments.

Money market deposits

These are fixed-interest term deposits of up to one year with banks and securities houses. They are also known as *time deposits* or *clean deposits*. They are not negotiable so cannot be liquidated before maturity. The interest rate on the deposit is fixed for the term and related to the London Interbank Offer Rate (Libor) of the same term. Interest and capital are paid on maturity.

Libor

The term LIBOR or 'Libor' comes from London Interbank Offered Rate, and is the interest rate at which one London bank offers funds to another London bank of acceptable credit quality in the form of a cash deposit. The rate is 'fixed' by the British Bankers Association at 1100 hours every business day morning (in practice the fix is usually about 20 minutes late) by taking the average of the rates supplied by member banks. The term Libid is the bank's 'bid' rate, that is the rate at which it pays for funds in the London market. The quote spread for a selected maturity is therefore the difference between Libor and Libid. The convention in London is to quote the two rates as Libor–Libid, thus matching the yield convention for other instruments. In some other markets the quote convention is reversed. Euribor is the interbank rate offered for euros as reported by the European Central Bank. Other money centres also have their rates fixed, so for example Stibor is the Stockholm banking rate, while pre-euro the Portuguese escudo rate fixing out of Lisbon was Lisbor.

The effective rate on a money market deposit is the annual equivalent interest rate for an instrument with a maturity of less than one year.

Certificates of deposit

Certificates of deposit (CDs) are receipts from banks for deposits that have been placed with them. They were first introduced in the sterling market in 1958. The deposits themselves carry a fixed rate of interest related to Libor and have a fixed term to maturity, so cannot be withdrawn before maturity. However the certificates themselves can be traded in a secondary market, that is, they are negotiable.[1] CDs are therefore very similar to negotiable money market deposits, although the yields are about 0.15% below the equivalent deposit rates because of the added benefit of liquidity. Most CDs issued are of between one and three months' maturity, although they do trade in maturities of one to five years. Interest is paid on maturity except for CDs lasting longer than one year, where interest is paid annually or occasionally semi-annually.

Banks, merchant banks and building societies issue CDs to raise funds to finance their business activities. A CD will have a stated interest rate and fixed maturity date, and can be issued in any denomination. On issue a CD is sold for face value, so the settlement proceeds of a CD on issue always equal its nominal value. The interest is paid, together with the face amount, on maturity. The interest rate is sometimes called the *coupon*, but unless the CD is held to maturity this will not equal the yield, which is of course the current rate available in the market and varies over time. In the United States CDs are available in smaller denomination amounts to retail investors.[2] The largest group of CD investors however are banks themselves, money market funds, corporates and local authority treasurers.

Unlike coupons on bonds, which are paid in rounded amounts, CD coupons are calculated to the exact day.

CD yields

The coupon quoted on a CD is a function of the credit quality of the issuing bank, and its expected liquidity level in the market, and of course the maturity of the CD, as this will be considered relative to the money market yield curve. As CDs are issued by banks as part of their short-term funding and liquidity requirement, issue volumes are driven by the demand for bank loans and the availability of alternative sources of funds for bank customers. The credit quality of the issuing bank is the primary consideration however; in the sterling market the lowest yield is paid by 'clearer' CDs, which are CDs issued by the clearing banks such as RBS NatWest, HSBC and Barclays plc. In the US market 'prime' CDs, issued by highly rated domestic banks, trade at a lower yield than non-prime CDs. In both markets CDs issued by foreign banks such as French or Japanese banks will trade at higher yields.

[1] A small number of CDs are non-negotiable.
[2] This was first introduced by Merrill Lynch in 1982.

Euro-CDs, which are CDs issued in a different currency from the home currency, also trade at higher yields, in the United States because of reserve and deposit insurance restrictions.

If the current market price of the CD including accrued interest is P and the current quoted yield is r, the yield can be calculated given the price, using (3.2).

$$r = \left\{ \frac{M}{P} \left[1 + C \left(\frac{N_{im}}{B} \right) \right] - 1 \right\} \times \left(\frac{B}{N_{sm}} \right) \tag{3.2}$$

The price can be calculated given the yield using (3.3).

$$\begin{aligned} P &= M \times \left[1 + C \left(\frac{N_{im}}{B} \right) \right] / \left[1 + r \left(\frac{N_{sm}}{B} \right) \right] \\ &= F / \left[1 + r \left(\frac{N_{sm}}{B} \right) \right] \end{aligned} \tag{3.3}$$

where

C is the quoted coupon on the CD
M is the face value of the CD
B is the year day-basis (365 or 360)
F is the maturity value of the CD
N_{im} is the number of days between issue and maturity
N_{sm} is the number of days between settlement and maturity
N_{is} is the number of days between issue and settlement.

After issue a CD can be traded in the secondary market. The secondary market in CDs in the UK is very liquid, and CDs will trade at the rate prevalent at the time, which will invariably be different from the coupon rate on the CD at issue. When a CD is traded in the secondary market, the settlement proceeds will need to take into account interest that has accrued on the paper and the different rate at which the CD has now been dealt. The formula for calculating the settlement figure is given at (3.4), which applies to the sterling market and its 365-day count basis.

$$\text{Proceeds} = \frac{M \times \text{Tenor} \times C \times 100 + 36500}{\text{Days remaining} \times r \times 100 + 36500} \tag{3.4}$$

The *tenor* of a CD is the life of the CD in days, while *days remaining* is the number of days left to maturity from the time of trade.

The return on holding a CD is given by (3.5).

$$\text{Return} = \left[\frac{\left(1 + \text{purchase yield} \times \dfrac{\text{days from purchase to maturity}}{B} \right)}{1 + \text{sale yield} \times \dfrac{\text{days from sale to maturity}}{B}} - 1 \right] \times \frac{B}{\text{days held}} \tag{3.5}$$

Securities quoted on a discount basis

The remaining money market instruments are all quoted on a *discount* basis, and so are known as 'discount' instruments. This means that they are issued on a discount to face value, and are redeemed on maturity at face value. Treasury bills, bills of exchange, bankers acceptances and commercial paper are examples of money market securities that are quoted on this basis: that is, they are sold on the basis of a discount to par. The difference between the price paid at the time of purchase and the redemption value (par) is the interest earned by the holder of the paper. Explicit interest is not paid on discount instruments, rather interest is reflected implicitly in the difference between the discounted issue price and the par value received at maturity. Note that in some markets CP is quoted on a yield basis, but not in the UK or in the United States where it is a discount instrument.

Treasury bills

Treasury bills or T-bills are government 'IOUs' of short duration, often three-month maturity. For example if a bill is issued on 10 January it will mature on 10 April. Bills of one-month and six-month maturity are also issued, but only rarely in the UK market. On maturity the holder of a T-bill receives the par value of the bill by presenting it to the Central Bank. In the UK most such bills are denominated in sterling but issues are also made in euros. In a capital market, T-bill yields are regarded as the *risk-free* yield, as they represent the yield from short-term government debt. In emerging markets they are often the most liquid instruments available for investors.

A sterling T-bill with £10 million face value issued for 91 days will be redeemed on maturity at £10 million. If the three-month yield at the time of issue is 5.25%, the price of the bill at issue is:

$$P = \frac{10\text{ m}}{\left(1 + 0.0525 \times \frac{91}{365}\right)} = £9{,}870{,}800.69$$

In the UK and US markets the interest rate on discount instruments is quoted as a *discount rate* rather than a yield. This is the amount of discount expressed as an annualised percentage of the face value, and not as a percentage of the original amount paid. By definition the discount rate is always lower than the corresponding yield. If the discount rate on a bill is d, then the amount of discount is given by (3.6):

$$d_{value} = M \times d \times \frac{n}{B} \tag{3.6}$$

The price P paid for the bill is the face value minus the discount amount, given by:

$$\text{P} = 100 \times \left[1 - \frac{d\left(\frac{N_{sm}}{365}\right)}{100}\right] \tag{3.7}$$

If we know the yield on the bill then we can calculate its price at issue by using the simple present value formula, as shown in (3.8):

$$P = M / \left[1 + r \left(\frac{N_{sm}}{365} \right) \right] \tag{3.8}$$

The discount rate d for T-bills is calculated using (3.9):

$$d = (1 - P) \times \frac{B}{n} \tag{3.9}$$

The relationship between discount rate and true yield is given by:

$$d = \frac{r}{\left(1 + r \times \frac{n}{B} \right)} \tag{3.10}$$

$$r = \frac{d}{\left(1 - d \times \frac{n}{B} \right)}$$

If a T-bill is traded in the secondary market, the settlement proceeds from the trade are calculated using (3.11):

$$\text{Proceeds} = M - \left(\frac{M \times \text{days remaining} \times d}{B \times 100} \right) \tag{3.11}$$

Bankers acceptances

A bankers acceptance is a written promise issued by a borrower to a bank to repay borrowed funds. The lending bank lends funds and in return accepts the bankers acceptance. The acceptance is negotiable and can be sold in the secondary market. The investor who buys the acceptance can collect the loan on the day that repayment is due. If the borrower defaults, the investor has legal recourse to the bank that made the first acceptance. Bankers acceptances are also known as *bills of exchange, bank bills, trade bills* or *commercial bills*.

Essentially bankers acceptances are instruments created to facilitate commercial trade transactions. The instrument is called a *bankers acceptance* because a bank accepts the ultimate responsibility to repay the loan to its holder. The use of bankers acceptances to finance commercial transactions is known as *acceptance financing*. The transactions for which acceptances are created include import and export of goods, the storage and shipping of goods between two overseas countries, where neither the importer nor the exporter is based in the home country,[3] and the storage and shipping of goods between two entities based at home. Acceptances are discount instruments and are purchased by banks, local authorities and

[3] A bankers acceptance created to finance such a transaction is known as a *third-party acceptance*.

money market investment funds. The rate that a bank charges a customer for issuing a bankers acceptance is a function of the rate at which the bank thinks it will be able to sell it in the secondary market. A commission is added to this rate. For ineligible bankers acceptances (see below) the issuing bank will add an amount to offset the cost of the additional reserve requirements.

Eligible bankers acceptance

An accepting bank that chooses to retain a bankers acceptance in its portfolio may be able to use it as collateral for a loan obtained from the central bank during open market operations, for example the Bank of England in the UK and the Fed in the United States. Not all acceptances are eligible to be used as collateral in this way, as they must meet certain criteria set by the central bank. The main requirement for eligibility is that the acceptance must be within a certain maturity band (a maximum of six months in the United States and three months in the UK), and that it must have been created to finance a self-liquidating commercial transaction. In the United States eligibility is also important because the Fed imposes a reserve requirement on funds raised via bankers acceptances that are ineligible. Bankers acceptances sold by an accepting bank are potential liabilities of the bank, but the Fed imposes a limit on the amount of eligible bankers acceptances that a bank may issue. Bills eligible for deposit at a central bank enjoy a finer rate than ineligible bills, and also act as a benchmark for prices in the secondary market.

Commercial paper

Commercial paper (CP) is a short-term money market funding instrument issued by corporates. In the UK and United States it is a discount instrument. Companies' short-term capital and *working capital* requirements are usually sourced directly from banks, in the form of bank loans. CP is an alternative short-term funding instrument, which is available to corporates that have a sufficiently strong credit rating. CP is an unsecured promissory note. The issuer of the note promises to pay its holder a specified amount on a specified maturity date. CP normally has a zero coupon and trades at a *discount* to its face value. The discount represents interest to the investor in the period to maturity. CP is typically issued in bearer form, although some issues are in registered form.

In the London market CP was not introduced until the mid-1980s. In the United States however, the market was developed in the late nineteenth century, and as early as 1922 there were 2200 issuers of CP with US$700 million outstanding. After its introduction in the UK in 1986, CP was subsequently issued in other European countries.

Originally the CP market was restricted to borrowers with high credit rating, and although lower-rated borrowers do now issue CP, sometimes by obtaining credit enhancements or setting up collateral arrangements, issuance in the market is still dominated by highly rated companies. The majority of issues are very short-term, from 30 to 90 days to maturity; it is extremely rare to observe paper with a maturity of more than 270 days or nine months. This is because of regulatory requirements

US Treasury bills

The Treasury bill market in the United States is one of the most liquid and transparent debt markets in the world. Consequently the bid–offer spread on the bills is very narrow. The Treasury issues bills at a weekly auction each Monday, made up of 91-day and 182-day bills. Every fourth week the Treasury also issues 52-week bills. As a result there are large numbers of Treasury bills outstanding at any one time. The interest earned on Treasury bills is not liable to state and local income taxes.

Federal funds

Commercial banks in the United States are required to keep reserves on deposit at the Fed. Banks with reserves in excess of required reserves can lend these funds to other banks, and these interbank loans are called *federal funds* or *fed funds* and are usually overnight loans. Through the fed funds market, commercial banks with excess funds are able to lend to banks that are short of reserves, thus facilitating liquidity. The transactions are very large denominations, and are lent at the *fed funds rate*, which is a very volatile interest rate because it fluctuates with market shortages.

Prime rate

The *prime interest rate* in the United States is often said to represent the rate at which commercial banks lend to their most creditworthy customers. In practice many loans are made at rates below the prime rate, so the prime rate is not the best rate at which highly rated firms may borrow. Nevertheless the prime rate is a benchmark indicator of the level of US money market rates, and is often used as a reference rate for floating-rate instruments. As the market for bank loans is highly competitive, all commercial banks quote a single prime rate, and the rate for all banks changes simultaneously.

in the United States,[4] which state that debt instruments with a maturity of less than 270 days need not be registered. Companies therefore issue CP with a maturity lower than nine months and so avoid the administration costs associated with registering issues with the Securities and Exchange Commission (SEC).

Table 3.1 is a comparison of US and Eurocommercial CP issues.

There are two major markets, the domestic US dollar market and the Eurocommercial paper market. CP markets are wholesale markets, and transactions are typically very large. In the United States over a third of all CP is purchased by

[4]This is the Securities Act of 1933. Registration is with the Securities and Exchange Commission.

Table 3.1 Comparison of US CP and Eurocommercial CP

	US CP	*Eurocommercial CP*
Currency	US dollar	Any Euro currency
Maturity	1–270 days	2–365 days
Typical maturity	30–60 days	30–90 days
Interest	Zero coupon, issued at discount	Zero-coupon, issued at discount
Quotation	On a discount rate basis	On a yield basis
Settlement	T + 0	T + 2
Registration	Bearer form	Bearer form
Negotiable	Yes	Yes

money market unit trusts, known as mutual funds; other investors include pension fund managers, retail or commercial banks, local authorities and corporate treasurers.

Although there is a secondary market in CP, very little trading activity takes place since investors generally hold CP until maturity. This is to be expected because investors purchase CP that matches their specific maturity requirement. When an investor does wish to sell paper, it can be sold back to the dealer, or when the issuer has placed the paper directly in the market (and not via an investment bank), it can be sold back to the issuer.

Commercial paper programmes

The issuers of CP are often divided into two categories of company, banking and financial institutions and non-financial companies. The majority of CP issues are by financial companies. Financial companies include not only banks but the financing arms of corporates such as General Motors and Ford Motor Credit. Most of the issuers have strong credit ratings, but lower-rated borrowers have tapped the market, often after arranging credit support from a higher-rated company, such as a *letter of credit* from a bank, or by arranging collateral for the issue in the form of high-quality assets such as Treasury bonds. CP issued with credit support is known as *credit-supported commercial paper*, while paper backed with assets is known naturally enough, as *asset-backed commercial paper*. Paper that is backed by a bank letter of credit is termed *LOC paper*. Although banks charge a fee for issuing letters of credit, borrowers are often happy to arrange for this, since by so doing they are able to tap the CP market. The yield paid on an issue of CP will be lower than a commercial bank loan.

Although CP is a short-dated security, typically of three to six-month maturity, it is issued within a longer-term programme, usually for three to five years for euro paper. US CP programmes are often open-ended. For example a company might arrange a five-year CP programme with a limit of US$5 billion. Once the programme is established the company can issue CP up to this amount, say for maturities of 30 or 60 days. The programme is continuous and new CP can be

issued at any time, daily if required. The total amount in issue cannot exceed the limit set for the programme. A CP programme can be used by a company to manage its short-term liquidity, that is, its working capital requirements. New paper can be issued whenever a need for cash arises, and for an appropriate maturity.

Issuers often roll over their funding and use funds from a new issue of CP to redeem a maturing issue. There is a risk that an issuer might be unable to roll over the paper where there is a lack of investor interest in the new issue. To provide protection against this risk, issuers often arrange a stand-by line of credit from a bank, normally for all of the CP programme, to draw against in the event that it cannot place a new issue.

There are two methods by which CP is issued, known as *direct-issued* or *direct paper* and *dealer-issued* or *dealer paper*. Direct paper is sold by the issuing firm directly to investors, and no agent bank or securities house is involved. It is common for financial companies to issue CP directly to their customers, often because they have continuous programmes and constantly roll over their paper. It is therefore cost-effective for them to have their own sales arm and sell their CP direct. The treasury arms of certain non-financial companies also issue direct paper; this includes for example British Airways plc corporate treasury, which runs a continuous direct CP programme, used to provide short-term working capital for the company. Dealer paper is paper that is sold using a banking or securities house intermediary. In the United States, dealer CP is effectively dominated by investment banks, as retail (commercial) banks were until recently forbidden from underwriting commercial paper. This restriction has since been removed and now both investment banks and commercial paper underwrite dealer paper.

Commercial paper yields

Commercial paper is a discount instrument. There have been issues of coupon CP in the euro market, but this is unusual. Thus CP is sold at a discount to its maturity value, and the difference between this maturity value and the purchase price is the interest earned by the investor. The CP day-count base is 360 days in the US and euro markets, and 365 days in the UK. The paper is quoted on a discount yield basis, in the same manner as Treasury bills. The yield on CP follows that of other money market instruments and is a function of the short-dated yield curve. The yield on CP is higher than the T-bill rate; this is because of the credit risk that the investor is exposed to when holding CP, for tax reasons (in certain jurisdictions interest earned on T-bills is exempt from income tax) and because of the lower level of liquidity available in the CP market. CP also pays a higher yield than certificates of deposit (CD), because of the lower liquidity of the CP market.

Although CP is a discount instrument and trades as such in the United States and UK, euro currency Eurocommercial paper trades on a yield basis, similar to a

CD. The expressions below illustrate the relationship between true yield and discount rate.

$$P = \frac{M}{1 + r' \dfrac{\text{days}}{\text{year}}} \tag{3.12}$$

$$rd = \frac{r}{1 + r' \dfrac{\text{days}}{\text{year}}} \tag{3.13}$$

$$r = \frac{rd}{1 + r' \dfrac{\text{days}}{\text{year}}} \tag{3.14}$$

where M is the face value of the instrument, rd is the discount rate and r the true yield.

Asset-backed commercial paper

Introduction

The rise in securitisation has led to the growth of short-term instruments backed by the cash flows from other assets, known as *asset-backed commercial paper* (ABCP). *Securitisation* is the practice of using the cash flows from a specified asset, such as residential mortgages, car loans or commercial bank loans, as backing for an issue of bonds. The assets themselves are transferred from the original owner (the *originator*) to a specially created legal entity known as a *special purpose vehicle* (SPV), so as to make them separate and bankruptcy-remote from the originator. In the meantime, the originator is able to benefit from capital market financing, often charged at a lower rate of interest than that earned by the originator on its assets. Securitised products are not money market instruments, and although ABCP is, most textbooks treat ABCP as part of the structured products market rather than as a money market product.

Generally securitisation is used as a funding instrument by companies for three main reasons. First, it offers lower-cost funding than traditional bank loan or bond financing. Second, it is a mechanism by which assets such as corporate loans or mortgages can be removed from the balance sheet, thus improving the lender's return on assets or return on equity ratios; and third, it increases a borrower's funding options. When entering into securitisation, an entity may issue term securities against assets into the public or private market, or it may issue commercial paper via a special vehicle known as a *conduit*. These conduits are usually sponsored by commercial banks.

Entities usually access the commercial paper market in order to secure permanent financing, rolling over individual issues as part of a longer-term programme

and using interest-rate swaps to arrange a fixed rate if required. Conventional CP issues are typically supported by a line of credit from a commercial bank, so this form of financing is in effect a form of bank funding. Issuing ABCP enables an originator to benefit from money market financing to which it might otherwise not have access because its credit rating is not sufficiently strong. A bank may also issue ABCP for balance sheet or funding reasons. ABCP trades exactly as conventional CP. The administration and legal treatment is more onerous, however, because of the need to establish the CP trust structure and issuing SPV. The servicing of an ABCP programme follows that of conventional CP and is carried out by the same entities, such as the 'trust' arms of banks such as JPMorgan Chase, Deutsche Bank and Bank of New York.

The example on page 40 details a hypothetical ABCP issue and typical structure.

Basic characteristics

Asset-backed CP programmes are invariably issued out of specially-incorporated legal entities (the SPV, sometimes called the SPC, for special purpose corporation). These conduits are typically established by commercial banks and finance companies to enable them to access Libor-based funding, at close to Libor, and to obtain regulatory capital relief. This can be done for the bank or a customer.

An ABCP conduit has the following features:

- It is a bankruptcy-remote legal entity that issues commercial paper to finance a purchase of assets from a seller of assets.
- The interest on the CP issued by the conduit, and its principal on maturity, will be paid out of the receipts on the assets purchased by the conduit.
- Conduits have also been set up to exploit credit arbitrage opportunities, such as raising finance at Libor to invest in high-quality assets such as investment-grade rated structured finance securities that pay above Libor.

The assets that can be funded via a conduit programme are many and varied. To date they have included:

- trade receivables and equipment lease receivables
- credit card receivables
- auto loans and leases
- corporate loans, franchise loans, mortgage loans
- real-estate leases
- investment-grade rated structured finance bonds such as ABS, MBS and CDO notes
- future (expected) cash flows.

Conduits are classified into a 'programme type', which refers to the make-up of the underlying asset portfolio. This can be single-seller or multi-seller, which indicates how many institutions or entities are selling assets to the conduit. They are also designated as funding or securities credit arbitrage vehicles. A special class of

conduit known as a structured investment vehicle (SIV, sometimes called a special investment vehicle) exists, that issues both CP and medium-term notes (MTNs), which are usually credit arbitrage vehicles.

Credit enhancement and liquidity support

To make the issue of liabilities from a conduit more appealing to investors (or to secure a particular credit rating), a programme sponsor will usually arrange some form of credit enhancement and/or back-up borrowing facility. Generally two types of credit enhancement are used, either 'pool-specific' or 'programme-wide' enhancement. The first arrangement will cover only losses on a specific named part of the asset pool, and cannot be used to cover losses in any other part of the asset pool. Programme-wide credit enhancement is a fungible layer of credit protection that can be drawn on to cover losses from the start or if any pool-specific facility has been used up.

Pool-specific credit enhancement instruments include the following:

- over-collateralisation, where the nominal value of the underlying assets exceeds that of the issued paper
- surety bond: a guarantee of repayment from a sponsor or other bank
- letter of credit: a standby facility that the issuer can use to draw funds from
- irrevocable loan facility
- excess cash, invested in eligible instruments such as T-bills.

The size of a pool-specific credit enhancement facility is quoted as a fixed percentage of the asset pool. Programme-wide credit enhancement is in the same form as pool-specific enhancement, and acts as a second layer of credit protection. It may be provided by a third party such as a commercial bank as well as by the sponsor.

Liquidity support is separate from credit enhancement. While credit enhancement facilities cover losses due to asset default, liquidity providers undertake to make available funds should these be required for reasons other than asset default. A liquidity line is drawn on, if required, to ensure timely repayment of maturing CP. This might occur because of market disruption (such that the issuer could not place new CP), an inability of the issuer to roll maturing CP, or because of asset and liability mismatches. This last is the least serious situation, and reflects that in many cases long-dated assets are used to back short-dated liabilities, and cash flow dates often do not match. The availability of a liquidity arrangement provides comfort to investors that CP will be repaid in full and on time, and is usually arranged with a commercial bank. It is usually provided as a loan agreement, of an amount equal to 100% of the face amount of CP issued, under which the liquidity provider agrees to lend funds to the conduit as required. The security for the liquidity line comes from the underlying assets.

Figure 3.1 (overleaf) illustrates a typical ABC structure issuing to the USCP and ECP markets and Figure 3.2 shows a multi-seller conduit set up to issue in the ECP market.

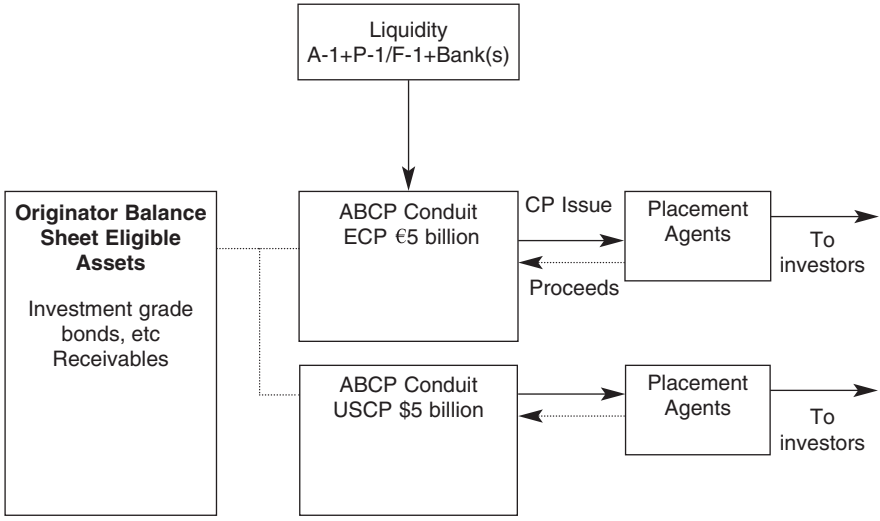


Figure 3.1 Single-seller ABCP conduit

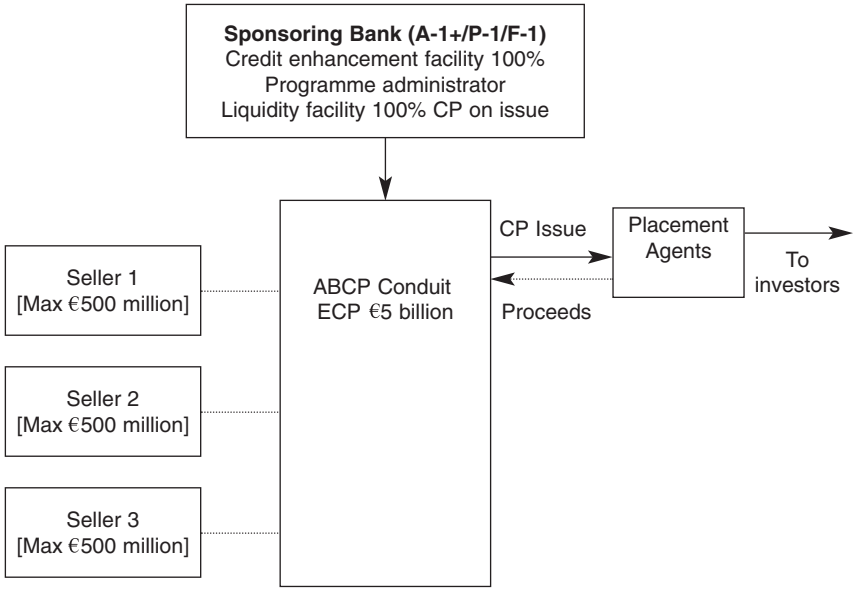


Figure 3.2 Multi-seller Euro ABCP conduit

Example 3.1

1. A 60-day CP note has a nominal value of £100,000. It is issued at a discount of 7.5% per annum. The discount is calculated as:

$$Dis = \frac{£100,000(0.075 \times 60)}{365}$$

$$= £1,232.88$$

The issue price for the CP is therefore £100,000 − £1,232, or £98,768.

The money market yield on this note at the time of issue is:

$$\left(\frac{365 \times 0.075}{365 - (0.075 \times 60)}\right) \times 100\% = 7.954\%$$

Another way to calculate this yield is to measure the capital gain (the discount) as a percentage of the CP's cost, and convert this from a 60-day yield to a one-year (365-day) yield, as shown below.

$$r = \frac{1,232}{98,768} \times \frac{365}{60} \times 100\%$$
$$= 7.588\%$$

2. ABC plc wishes to issue CP with 90 days to maturity. The investment bank managing the issue advises that the discount rate should be 9.5%. What should the issue price be, and what is the money market yield for investors?

$$Dis = \frac{100,000\ (0.095 \times 90)}{365}$$
$$= 2.342$$

The issue price will be 97.658.

The yield to investors will be:

$$\frac{2.342}{97.658} \times \frac{365}{90} \times 100\% = 9.725\%$$

Foreign exchange

The price quotation for currencies generally follows the ISO convention, which is also used by the SWIFT and Reuters dealing systems, and is the three-letter code used to identify a currency, such as USD for US dollar and GBP for sterling. The rate convention is to quote everything in terms of one unit of the US dollar, so that the dollar and Swiss franc rate is quoted as USD/CHF, and is the number of

Swiss francs to one US dollar. The exception is for sterling, which is quoted as GBP/USD and is the number of US dollars to the pound. The rate for euros has been quoted both ways round, for example EUR/USD although some banks, for example RBS Financial Markets in the UK, quote euros to the pound, that is GBP/EUR.

Spot exchange rates

A *spot* FX trade is an outright purchase or sale of one currency against another currency, with delivery two working days after the trade date. Non-working days do not count, so a trade on a Friday is settled on the following Tuesday. There are some exceptions to this, for example trades of US dollars against Canadian dollars are settled the next working day. Note that for some currencies, generally in the Middle East, markets are closed on Friday but open on Saturday. A settlement date that falls on a public holiday in the country of one of the two currencies is delayed for settlement by that day.

An FX transaction is possible between any two currencies. However to reduce the number of quotes that need to be made, the market generally quotes only against the US dollar or occasionally sterling or the euro, so that the exchange rate between two non-dollar currencies is calculated from the rate for each currency against the dollar. The resulting exchange rate is known as the *cross-rate*. Cross-rates themselves are also traded between banks in addition to dollar-based rates. This is usually because the relationship between two rates is closer than that of either against the dollar. For example the Swiss franc moves more closely in line with the euro than against the dollar, so in practice one observes that the dollar/Swiss franc rate is more a function of the euro/franc rate.

The spot FX quote is a two-way bid-offer price, just as in the bond and money markets, and indicates the rate at which a bank is prepared to buy the base currency against the variable currency; this is the 'bid' for the variable currency, so is the lower rate. The other side of the quote is the rate at which the bank is prepared to sell the base currency against the variable currency. For example a quote of 1.6245–1.6255 for GBP/USD means that the bank is prepared to buy sterling for US$1.6245, and to sell sterling for US$1.6255. The convention in the FX market is uniform across countries, unlike the money markets. Although the money market convention for bid–offer quotes is for example, 5.5–5.25%, meaning that the 'bid' for paper – the rate at which the bank will lend funds, say in the CD market – is the higher rate and always on the left, this convention is reversed in certain countries. In the FX markets the convention is always the same one just described.

The difference between the two sides in a quote is the bank's dealing spread. Rates are quoted to 1/100th of a cent, known as a *pip*. In the quote above, the spread is 10 pips. However this amount is a function of the size of the quote number, so that the rate for USD/JPY at, say, 110.10–110.20 indicates a spread of 0.10 yen. Generally only the pips in the two rates are quoted, so that for example the quote above would be simply '45–55'. The 'big figure' is not quoted.

The derivation of cross-rates can be depicted in the following way. If we assume two exchange rates XXX/YYY and XXX/ZZZ, the cross-rates are:

YYY/ZZZ = XXX/ZZZ ÷ XXX/YYY
ZZZ/YYY = XXX/YYY ÷ XXX/ZZZ

Given two exchange rates YYY/XXX and XXX/ZZZ, the cross-rates are:

YYY/ZZZ = YYY/XXX ÷ XXX/ZZZ
ZZZ/YYY = 1 ÷ (YYY/XXX ÷ XXX/ZZZ)

Forward exchange rates

Forward outright

The spot exchange rate is the rate for immediate delivery (notwithstanding that actual delivery is two days forward). A *forward contract* or simply *forward* is an outright purchase or sale of one currency in exchange for another currency for settlement on a specified date at some point in the future. The exchange rate is quoted in the same way as the spot rate, with the bank buying the base currency on the bid side and selling it on the offered side. In some emerging markets no liquid forward

Example 3.2: Exchange cross-rates

Consider the following two spot rates:

EUR/USD	1.0566–1.0571
AUD/USD	0.7034–0.7039

The EUR/USD dealer buys euros and sells dollars at 1.0566 (the left side), while the AUD/USD dealer sells Australian dollars and buys US dollars at 0.7039 (the right side). To calculate the rate at which the bank buys euros and sells Australian dollars, we need:

1.0566/0.7039 = 1.4997

This is the rate at which the bank buys euros and sells Australian dollars. In the same way the rate at which the bank sells euros and buys Australian dollars is given by:

1.0571/0.7034 or 1.5028

Therefore the spot EUR/AUD rate is 1.4997–1.5028.

market exists, so forwards are settled in cash against the spot rate on the maturity date. These *non-deliverable forwards* are considered at the end of this section.

Although some commentators have stated that the forward rate may be seen as the market's view of where the spot rate will be on the maturity date of the forward transaction, this is incorrect. A forward rate is calculated on the current interest rates of the two currencies involved, and the principle of no-arbitrage pricing ensures that there is no profit to be gained from simultaneous (and opposite) dealing in spot and forward. Consider the following strategy:

- Borrow US dollars for six months starting from the spot value date.
- Sell dollars and buy sterling for value spot.
- Deposit the long sterling position for six months from the spot value date.
- Sell forward today the sterling principal and interest which mature in six months time into dollars.

The market will adjust the forward price so that the two initial transactions if carried out simultaneously will generate a zero profit/loss. The forward rates quoted in the trade will be calculated on the six months' deposit rates for dollars and sterling; in general the calculation of a forward rate is given in equation (3.15).

$$Fwd = Spot \times \frac{\left(1 + \text{variable currency deposit rate} \times \frac{days}{B}\right)}{\left(1 + \text{variable currency deposit rate} \times \frac{days}{B}\right)} \tag{3.15}$$

The year day-count base B will be either 365 or 360 depending on the convention for the currency in question.

Example 3.3: Forward rate

90-day GBP deposit rate: 5.75%
90-day USD deposit rate: 6.15%
Spot GBP/USD rate: 1.6315 (mid-rate)

The forward rate is given by:

$$1.6315 \times \frac{\left(1 + 0.0575 \times \frac{90}{365}\right)}{\left(1 + 0.0575 \times \frac{90}{360}\right)} = 1.6296$$

Therefore to deal forward the GBP/USD mid-rate is 1.6296, so in effect £1 buys US$1.6296 in three months time as opposed to US$1.6315 today. Under different circumstances sterling may be worth more in the future than at the spot date.

Forward swaps

The calculation given as Example 3.3 illustrates how a forward rate is calculated and quoted in theory. In practice as spot rates change rapidly, often many times even in one minute, it would be tedious to keep recalculating the forward rate so often. Therefore banks quote a forward spread over the spot rate, which can then be added or subtracted to the spot rate as it changes. This spread is known as the *swap points*. An approximate value for the number of swap points is given in (3.16).

$$\text{Forward swap} \approx \text{spot} \times \text{deposit rate differential} \times \frac{\text{days}}{B} \tag{3.16}$$

The approximation is not accurate enough for forwards maturing more than 30 days from now, in which case another equation must be used. This is given as (3.17). It is also possible to calculate an approximate deposit rate differential from the swap points by rearranging (3.16).

$$\text{Forward swap} = \frac{\text{spot} \times \left(vc \text{ depo rate} \times \frac{\text{days}}{B} - bc \text{ depo rate} \times \frac{\text{days}}{B} \right)}{\left(1 + bc \text{ depo rate} \times \frac{\text{days}}{B} \right)} \tag{3.17}$$

where *vc* is variable currency and *bc* is base currency.

Example 3.4: Forward swap points

Spot EUR/USD:	1.0566–1.0571
Forward swap:	0.0125–0.0130
Forward outright:	1.0691–1.0701

The forward outright is the spot price + the swap points, so in this case,

1.0691 = 1.0566 + 0.0125
1.0701 = 1.0571 + 0.0130

Spot EUR/USD rate:	0.9501
31-day EUR rate:	3.15%
31-day USD rate:	5.95%

$$\text{Forward swap} = 0.9501 \times \left(0.0595 \times \frac{31}{365} - 0.0315 \times \frac{31}{360} \right) = 0.0024$$
$$\left(1 + 0.0315 \times \frac{31}{360} \right)$$

or +24 points.

The swap points are quoted as two-way prices in the same way as spot rates. In practice a middle spot price is used and then the forward swap spread around the spot quote. The difference between the interest rates of the two currencies will determine the magnitude of the swap points and whether they are added or subtracted from the spot rate. When the swap points are positive and the forwards trader applies a bid–offer spread to quote a two-way price, the left-hand side of the quote is smaller than the right-hand side as usual. When the swap points are negative, the trader must quote a 'more negative' number on the left and a 'more positive' number on the right-hand side. The 'minus' sign is not shown however, so that the left-hand side may appear to be the larger number. Basically when the swap price appears larger on the right, it means that it is negative and must be subtracted from the spot rate and not added.

Forwards traders are in fact interest rate traders rather than foreign exchange traders; although they will be left positions that arise from customer orders, in general they will manage their book based on their view of short-term deposit rates in the currencies they are trading. In general a forwards trader expecting the interest rate differential to move in favour of the base currency, for example, a rise in base currency rates or a fall in the variable currency rate, will 'buy and sell' the base currency. This is equivalent to borrowing the base currency and depositing in the variable currency. The relationship between interest rates and forward swaps means that banks can take advantage of different opportunities in different markets. Assume that a bank requires funding in one currency but is able to borrow in another currency at a relatively cheaper rate. It may wish to borrow in the second currency and use a forward contract to convert the borrowing to the first currency. It will do this if the all-in cost of borrowing is less than the cost of borrowing directly in the first currency.

Forward cross-rates

A forward cross-rate is calculated in the same way as spot cross-rates. The formulas given for spot cross-rates can be adapted to forward rates.

Forward-forwards

A forward-forward swap is a deal between two forward dates rather than from the spot date to a forward date; this is the same terminology and meaning as in the bond markets, where a forward or a forward-forward rate is the zero-coupon interest rate between two points both beginning in the future. In the foreign exchange market, an example would be a contract to sell sterling three months forward and buy it back in six months' time. Here, the swap is for the three-month period between the three-month date and the six-month date. The reason a bank or corporate might do this is to hedge a forward exposure or because of a particular view it has on forward rates, in effect deposit rates.

Example 3.5: Forward-forward contract

GBP/USD spot rate:	1.6315–20
3-month swap:	45–41
6-month swap:	135–125

If a bank wished to sell GBP three months forward and buy them back six months forward, this is identical to undertaking one swap to buy GBP spot and sell GBP three months forward, and another to sell GBP spot and buy it six months forward. Swaps are always quoted as the quoting bank buying the base currency forward on the bid side, and selling the base currency forward on the offered side; the counterparty bank can 'buy and sell' GBP 'spot against three months' at a swap price of −45, with settlement rates of spot and (spot − 0.0045). It can 'sell and buy' GBP 'spot against six months' at the swap price of −125 with settlement rates of spot and (spot − 0.0125). It can therefore do both simultaneously, which implies a difference between the two forward prices of (−125) − (−45) = −90 points. Conversely the bank can 'buy and sell' GBP 'three months against six months' at a swap price of (−135) − (−41) or −94 points. The two-way price is therefore 94−90 (we ignore the negative signs).

Selected bibliography and references

Choudhry, M. *The Money Markets Handbook: A Practitioner's Guide*, Wiley Asia, 2005.
Roth, P. *Mastering Foreign Exchange and Money Markets*, FT Prentice Hall, 1996.
Walmsley, J. *The Foreign Exchange and Money Markets Guide*, 2nd edn, Wiley, 2000.

4
Fixed Income Securities I: The Bond Markets

Introduction

In most countries government expenditure exceeds the level of government income received through taxation. This shortfall is met by government borrowing, and bonds are issued to finance the government's debt. The core of any domestic capital market is usually the government bond market, which also forms the benchmark for all other borrowing. Government agencies also issue bonds, as do local governments or municipalities. Often (but not always) these bonds are virtually as secure as government bonds. Corporate borrowers issue bonds both to raise finance for major projects and to cover ongoing and operational expenses. Corporate finance is a mixture of debt and equity, and a specific capital project will often be financed by a mixture of both.

The debt capital markets exist because of the financing requirements of governments and corporates. The sources of capital are varied, but the total supply of funds in a market is made up of personal or household savings, business savings and increases in the overall money supply. However, the requirements of savers and borrowers differ significantly, in that savers have a short-term investment horizon while borrowers prefer to take a longer-term view. The 'constitutional weakness' of what would otherwise be unintermediated financial markets led, from an early stage, to the development of financial intermediaries.

The world bond market has increased in size more than 15 times in the last 30 years. As at the end of 2008 outstanding volume stood at over US$23 trillion.

The origin of the spectacular increase in the size of global financial markets was the rise in oil prices in the early 1970s. Higher oil prices stimulated the development of a sophisticated international banking system, as they resulted in large capital inflows to developed country banks from the oil-producing countries. A significant proportion of these capital flows were placed in *Eurodollar* deposits in major banks. The growing trade deficit and level of public borrowing in the United States also contributed. The last 20 years have seen tremendous growth in capital markets' volumes and trading. As capital controls were eased and exchange rates moved from fixed to floating, domestic capital markets became internationalised. Growth was assisted by the rapid advance in information technology and

the widespread use of financial engineering techniques. Today we would think nothing of dealing in virtually any liquid currency bond in financial centres around the world, often at the touch of a button. Global bond issues, underwritten by the subsidiaries of the same banks, are commonplace. The ease with which transactions can be undertaken has also contributed to a very competitive market in liquid currency assets.

Market participants

A wide range of participants are involved in the bond markets. We can group them broadly into borrowers and investors, plus the institutions and individuals who are part of the business of bond trading. Borrowers access the bond markets as part of their financing requirements; hence, borrowers can include sovereign governments, local authorities, public sector organisations and corporations. Virtually all businesses operate with a financing structure that is a mixture of debt and equity finance, and debt finance almost invariably contains a form of bond finance.

Intermediaries and banks

In its simplest form a financial intermediary is a broker or agent. Today we would classify the broker as someone who acts on behalf of the borrower or lender, buying or selling a bond as instructed. However, intermediaries originally acted between borrowers and lenders in placing funds as required. A broker would not simply on-lend funds that had been placed with it, but would accept deposits and make loans as required by its customers. This resulted in the first banks.

A *retail bank* deals mainly with the personal financial sector and small businesses, and in addition to loans and deposits also provides cash transmission services. A retail bank is required to maintain a minimum cash reserve, to meet potential withdrawals, but the remainder of its deposit base can be used to make loans. This does not mean that the total size of its loan book is restricted to what it has taken in deposits: loans can also be funded in the wholesale market.

An *investment bank* will deal with governments, corporates and institutional investors. Investment banks perform an agency role for their customers and are the primary vehicle through which a corporate will borrow funds in the bond markets. This is part of the bank's corporate finance function. It will also act as wholesaler in the bond markets, a function known as *market making*. The bond-issuing function of an investment bank, by which the bank will issue bonds on behalf of a customer and pass the funds raised to this customer, is known as *origination*. Investment banks will also carry out a range of other functions for institutional customers, including export finance, corporate advisory services and fund management. Other financial intermediaries will trade not on behalf of clients but for their own book. These include arbitrageurs and speculators. Usually such market participants form part of investment banks.

Institutional investors

We can group the main types of institutional investors according to the time horizon of their investment activity:

- *Short-term institutional investors*. These include banks and building societies, money market fund managers, central banks and the treasury desks of some types of corporates. Such bodies are driven by short-term investment views, often subject to close guidelines. Banks will have an additional requirement to maintain liquidity, often in fulfilment of regulatory authority rules, by holding a proportion of their assets in the form of short-term instruments that are easy to trade.
- *Long-term institutional investors*. Typically these types of investors include pension funds and life assurance companies. Their investment horizon is long-term, reflecting the nature of their liabilities. Often they will seek to match these liabilities by holding long-dated bonds.
- *Mixed horizon institutional investors*. This is possibly the largest category of investors and will include general insurance companies and most corporate bodies. Like banks and financial sector companies, they are also very active in the primary market, issuing bonds to finance their operations.

Market professionals

These players include the banks and specialist financial intermediaries mentioned above, firms that one would not automatically classify as 'investors' although they will also have an investment objective. Their time horizon will range from one day to the very long term. They include:

- proprietary trading desks of investment banks
- bond market makers in securities houses and banks providing a service to their customers
- inter-dealer brokers that provide an anonymous broking facility.

Proprietary traders will actively position themselves in the market in order to gain trading profit, for example in response to their view on where they think interest rate levels are headed. These participants will trade direct with other market professionals and investors, or via brokers.

Market makers or 'traders' (also called 'dealers' in the United States) are wholesalers in the bond markets; they make two-way prices in selected bonds. Firms will not necessarily be active market makers in all types of bonds; smaller firms often specialise in certain sectors. In a two-way quote the *bid price* is the price at which the market maker will buy stock, so it is the price the investor will receive when selling stock. The *offer price* or ask price is the price at which investors can buy stock from the market maker. As one might expect, the bid price is always higher than the offer price, and it is this spread that represents the theoretical profit to the market maker. The bid–offer spread set by the market maker is determined by

several factors, including supply and demand, and liquidity considerations for that particular stock, the trader's view on market direction and volatility, as well as that of the stock itself and the presence of any market intelligence. A large bid–offer spread reflects low liquidity in the stock, as well as low demand.

To facilitate a liquid market there also exist *inter-dealer brokers* (IDBs). These provide an anonymous broking facility so that market makers can trade in size at the keenest prices. Generally IDBs will post prices on their screens that have been provided by market makers on a no-names basis. The screens are available to other market makers (and in some markets to other participants as well). At any time IDB screen prices represent the latest market price and bid–offer spread. IDBs exist in government, agency, corporate and Eurobond markets.

Bonds by issuers

This section describes the main classes of bonds by type of borrower. On the public side we distinguish between *sovereign bonds* issued by national governments, *agency bonds* issued by public bodies and *municipal bonds* issued by local governments. On the private side we have the *corporate bonds* issued by corporations, and we further distinguish between *domestic* and *foreign bonds*, and *international bonds*, the latter constituting the large class of *Eurobonds*. Here we discuss the special characteristics of each of these types of bond.

Government bonds

The four major government bond issuers in the world are the euro-area countries, Japan, the United States and the United Kingdom.

Table 4.1 compares the features of the world's most important government bond markets. Note the minor variations in market practice with regard to the frequency of coupons, the day-count basis, benchmark bonds and so on. Most government bonds are issued by a standard auction process, where the price is gradually reduced until it meets a bid. The sale price varies for each successful bidder, depending on the bid price. Others use the so-called Dutch auction system. Under this system the securities are allocated to bidders starting with the highest bid. The price at which the final allocation is made becomes the price at which *all* securities are sold.

Table 4.2 shows the country yield curves at the time of writing, and a subset of these are graphed in Figure 4.1. Note the variability in yield curves between countries reflecting their varying economic conditions and risk profiles. While most have an upward-sloping (or normal) yield curve, two of the yield curves (UK and Australia) are quite flat. Discussion of the various theories explaining the shape of the yield curve can be found in Chapter 7.

In the US case, government securities are issued by the US Department of the Treasury and backed by the full faith and credit of the US government. These are called 'Treasury securities'. The Treasury market is the most active market in the world, thanks to the large volume of total debt and the large size of any single issue. The amount of outstanding marketable US Treasury securities is huge, with

Table 4.1 Government bond markets: characteristics of selected countries

Country	*Credit rating*	*Maturity range*	*Dealing*	*Benchmark bonds*	*Issuance*	*Coupon and day-count basis*
Australia	AAA	2–15 years	OTC dealer network	5, 10 years	Auction	Semi-annual, act/act
Canada	AAA	2–30 years	OTC dealer network	3, 5, 10 years	Auction, subscription	Semi-annual, act/act
France	AAA	BTAN: 1–7 years OAT: 10–30 years	OTC dealer network Bonds listed on Paris Stock Exchange	BTAN: 2 and 5 year OAT: 10 and 30 years	Dutch auction	BTAN: Semi–annual, act/act OAT: Annual, act/act
Germany	AAA	OBL: 2, 5 years BUND: 10, 30 years	OTC dealer network Listed on Stock Exchange	The most recent issue	Combination of Dutch auction and proportion of each issue allocated on fixed basis to institutions	Annual, act/act
Singapore	AAA	2–15 years	OTC dealer network	1, 5, 10 and 15 years	Auction	Semi-annual, act/act
South Africa	A	2–30 years	OTC dealer network Listed on Johannesburg SE	2, 7, 10 and 20 years	Auction	Semi–annual, act/365
Taiwan	AA−	2–30 years	OTC dealer network	2, 5, 10, 20 and 30 years	Auction	Annual, act/act
United Kingdom	AAA	2–35 years	OTC dealer network	5, 10, 30 years	Auction, subsequent issue by 'tap' subscription	Semi-annual, act/act
United States	AAA	2–20 years	OTC dealer network	2, 5, 10 years	Auction	Semi-annual, act/act

Source: Choudhry (2004).

Table 4.2 Country yield curves (as of 28 May 2009)

Tenor	*Australia*	*Canada*	*France*	*Germany*	*Singapore*	*South Africa*	*Taiwan*	*UK*	*US*
1	3.108	0.477	0.9345	0.8902	0.338	6.563	0.3927	0.6794	0.473
2	3.6441	1.2703	1.4108	1.4275	0.4475			1.1324	0.9621
3	4.1039	1.8271	1.9133	1.8492				1.9945	1.4832
4	4.501	2.2821	2.4773	2.3933		8.1031		2.5208	
5	4.7869	2.5572	2.8819	2.7228	1.1989		0.9678	2.6027	2.4169
7	5.2109	3.0444	3.446	3.2914	1.6813	8.202		2.8025	3.1769
10	5.3761	3.5707	4.011	3.6783	2.5459	8.721	1.5855	3.8222	3.6779
15	5.549	4.0182	4.372	4.2327	2.9422	8.845	2.2017	4.3751	
20			4.584	4.4605	3.1956	8.755	2.3132	4.5168	
30			4.579	4.4258		8.355	2.4473	4.6283	4.5935

Source: Bloomberg L.P.

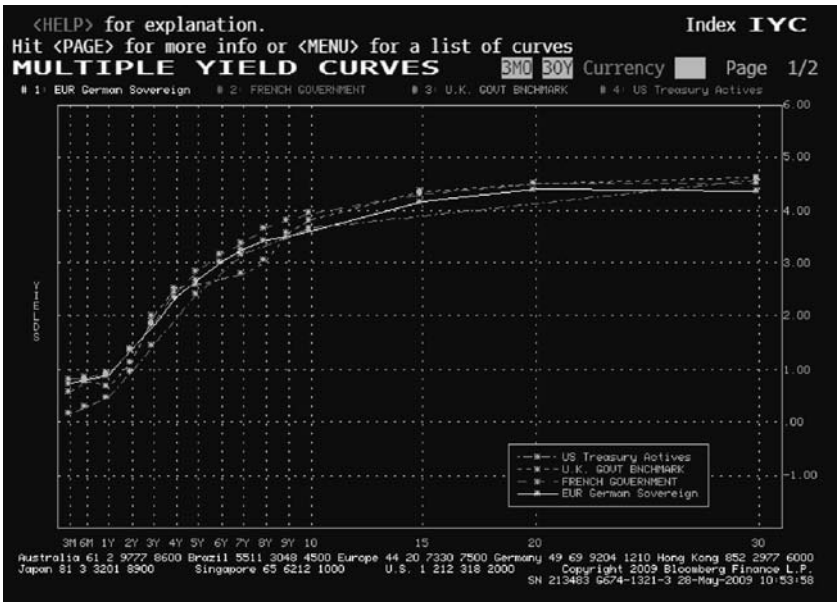


Figure 4.1 Bloomberg screen IYC showing yield curves for US, UK, French and German government bond markets, 28 May 2009
© Bloomberg L.P. Used with permission. All rights reserved. Visit www.bloomberg.com.

a value of US$4.2 trillion as of December 2008. The Treasury market is the most liquid debt market, that is, the one where pricing and trading are most efficient. The bid–offer spread is far lower than in the rest of the bond market. Recently issued Treasury securities are referred to as *on-the-run* securities, as opposed to *off-the-run* securities, which are old issued securities. Special mention must be made of *benchmark securities*, which are recognised as market indicators. There typically exists one such security on each of the following yield curve points: 2 years, 5 years, 10 years and 30 years. As they are over-liquid they trade richer than all of their direct neighbours.

Government bonds are traded on major exchanges as well as *over-the-counter*.[1] The New York Stock Exchange had over 700 government issues listed on it at the end of 2008, with a total par value of US$4.1 trillion.

US agency bonds

These are issued by different organisations, seven of which dominate the US market in terms of outstanding debt: the Federal National Mortgage Association (Fannie Mae), the Federal Home Loan Bank System (FHLBS), the Federal Home

[1]Generally OTC refers to trades that are not carried out on an exchange but directly between the counterparties.

Loan Mortgage Corporation (Freddie Mac), the Farm Credit System (FCS), the Student Loan Marketing Association (Sallie Mae), the Resolution Funding Corporation (REFCO) and the Tennessee Valley Authority (TVA). Agencies have at least two common features:

- *They were created to fulfil a public purpose*. For example in the United States, Fannie Mae and Freddie Mac aim to provide liquidity for the residential mortgage market. The FCS aims to support agricultural and rural lending. REFCO aims to provide financing to resolve thrift crisis.
- *The debt is not necessarily guaranteed by the government*. Hence it contains a credit premium. In fact in the United States, there are a few federally related institution securities, such as the Government National Mortgage Association (GNMA), and these are generally backed by the full faith and credit of the US government. There is no credit risk, but since they are relatively small issues they contain a liquidity premium.

Agencies are differently organised. For instance, Sallie Mae is owned by private-sector shareholders, and the FCS and the FHLBS are cooperatives owned by the members and borrowers. One sizeable agency, the Tennessee Valley Authority, is owned by the US government. Fannie Mae and Freddie Mae were taken over by the US government in 2008.

Municipal bonds

Municipal securities constitute the municipal market, that is, the market where state and local governments – counties, special districts, cities and towns – raise funds in order to finance projects for the public good such as schools, highways, hospitals, bridges and airports. Typically, bonds issued in this sector are exempt from federal income taxes, so this sector is referred to as the *tax-exempt sector*. There are two generic types of municipal bonds: *general obligation bonds* and *revenue bonds*. The former have principal and interest secured by the full faith and credit of the issuer and are usually supported by either the issuer's unlimited or limited taxing power. The latter have principal and interest secured by the revenues generated by the operating projects financed with the proceeds of the bond issue. Many of these bonds are issued by special authorities created for the purpose.

Corporate bonds

Corporate bonds are issued by entities belonging to the private sector. They represent what market participants call the credit market. In the corporate markets, bond issues usually have a stated term to maturity, although the term is often not fixed because of the addition of call or put features. The convention is for most corporate issues to be medium- or long-dated, and rarely to have a term greater than 20 years. In the US market prior to the Second World War it was once common for companies to issue bonds with maturities of 100 years or more, but this is now quite rare. Only the highest-rated companies find it possible to issue

bonds with terms to maturity greater than 30 years; during the 1990s such companies included Coca-Cola, Disney and British Gas.

Investors prefer to hold bonds with relatively short maturities because of the greater price volatility experienced in the markets since the 1970s, when high inflation and high interest rates were common. A shorter-dated bond has lower interest-rate risk and price volatility than a longer-dated bond. There is thus a conflict between investors, whose wish is to hold bonds of shorter maturities, and borrowers, who would like to fix their borrowing for as long a period as possible. Although certain institutional investors such as pension fund managers have an interest in holding 30-year bonds, it would be difficult for all but the largest, best-rated companies, to issue debt with a maturity greater than this.

Highly rated corporate borrowers are often able to issue bonds without indicating specifically how they will be redeemed. By implication, maturity proceeds will be financed out of the company's general operations or by the issue of another bond. However, borrowers with low ratings may make specific provisions for paying off a bond issue on its maturity date, to make their debt issue more palatable to investors. For instance, a ring-fenced sum of cash (called the *sinking fund*) may be put aside to form the proceeds used in the repayment of a fixed-term bond. A proportion of a bond issue is redeemed every year until the final year when the remaining outstanding amount is repaid. In most cases the issuer will pass the correct cash proceeds to the bond's trustee, who will use a lottery method to recall bonds representing the proportion of the total nominal value outstanding that is being repaid. The trustee usually publishes the serial numbers of bonds that are being recalled in a newspaper such as the *Wall Street Journal* or the *Financial Times*.

The price at which bonds are redeemed by a sinking fund is usually par. If a bond has been issued above par, the sinking fund may retire the bonds at the issue price and gradually decrease this each year until it reaches par.[2] Sinking funds reduce the credit risk applying to a bond issue, because they indicate to investors that provision has been made to repay the debt. However, there is a risk associated with them, in that at the time bonds are paid off they may be trading above par due to a decline in market interest rates. In this case investors will suffer a loss if it is their holding that is redeemed.

Bonds that are secured through a charge on fixed assets such as property or plant often have certain clauses in their offer documents that state that the issuer cannot dispose of the assets without making provision for redemption of the bonds, as this would weaken the collateral backing for the bond. These clauses are known as *release-of-property* and *substitution-of-property* clauses. Under these clauses, if property or plant is disposed, the issuer must use the proceeds (or part of the proceeds) to redeem bonds that are secured by the disposed assets. The price at which the bonds are retired under this provision is usually par, although a special redemption price other than par may be specified in the repayment clause.

[2]Another method by which bonds are repaid is that the issuer will purchase the required nominal value of the bonds in the open market; these are then delivered to the trustee, who cancels them.

A large number of corporate bonds, especially in the US market, have a *call provision*. Borrowers prefer this as it enables them to refinance debt at cheaper rates when market interest rates have fallen significantly below their level at the time of the bond issue. A call provision is a negative feature for investors, as bonds are only paid off if their price has risen above par. Although a call feature indicates an issuer's interest in paying off the bond, because they are not attractive for investors, callable bonds pay a higher yield than non-callable bonds of the same credit quality.

In general, callable bonds are not callable for the first five to ten years of their life, a feature that grants an element of protection for investors. Thereafter a bond is usually callable on set dates up to the final maturity date. In the US market another restriction is that of refunding redemption. This prohibits repayment of bonds within a set period after issue with funds obtained at a lower interest rate or through issue of bonds that rank with or ahead of the bond being redeemed. A bond with refunding protection during the first five to ten years of its life is not as attractive as a bond with absolute call protection. Bonds that are called are usually called at par, although it is common also for bonds to have a call schedule that states that they are redeemable at specified prices above par during the call period.

Corporate bonds are traded on exchanges and OTC. Outstanding volume as at the end of 2003 was US$8.1 trillion (see Choudhry, 2004). The corporate bond market varies in liquidity, depending on the currency and type of issuer of any particular bond. As in the case of sovereign bonds, liquidity is greater for recent issues. But corporate bonds in general are far less liquid than government bonds: they bear higher bid–ask spreads.

Eurobonds (international bonds)

In any market there is a primary distinction between *domestic* bonds and other bonds. Domestic bonds are issued by borrowers domiciled in the country of issue, and in the currency of the country of issue. Generally they trade only in their original market. A *Eurobond* is issued across national boundaries and can be in any currency, which is why it is also sometimes called an *international* bond. In fact, it is now more common for Eurobonds to be referred to as international bonds, to avoid confusion with 'euro bonds', which are bonds denominated in *euros*, the currency of 12 countries of the European Union (EU). As an issue of international bonds is not restricted in terms of currency or country, the borrower is not restricted as to its nationality either. There are also *foreign* bonds, which are domestic bonds issued by foreign borrowers. An example of a foreign bond is a *Bulldog*, which is a sterling bond issued for trading in the UK market by a foreign borrower. The equivalent foreign bonds in other countries include *Yankee* bonds (United States), *Samurai* bonds (Japan), *Alpine* bonds (Switzerland) and *Matador* bonds (Spain). There are detailed differences between these bonds, for example in the frequency of interest payments that each one makes and the way the interest payment is calculated. Some bonds, such as domestic bonds, pay their interest *net*, which means net of a withholding tax

such as income tax. Other bonds, including Eurobonds, make *gross* interest payments.

Nowhere has the increasing integration and globalisation of the world's capital markets been more evident than in the Eurobond market. It is an important source of funds for many banks and corporates, not to mention central governments. The Eurobond market continues to develop new structures in response to the varying demands and requirements of specific groups of investors. Often the Eurobond market is the only opening for certain types of government and corporate finance. Investors also look to the Eurobond market due to constraints in their domestic market, and Euro securities have been designed to reproduce the features of instruments that certain investors may be prohibited from investing in domestically. Other instruments are designed for investors in order to provide tax advantages. The traditional image of the Eurobond investor, the so-called 'Belgian dentist', has changed and the investor base is both varied and geographically dispersed.

The markets

A distinction is made between financial instruments of up to one year's maturity and instruments of over one year's maturity. Short-term instruments make up the *money market* while all other instruments are deemed to be part of the *capital market*. There is also a distinction made between the *primary market* and the *secondary market*. A new issue of bonds made by an investment bank on behalf of its client is made in the primary market. Such an issue can be a *public offer*, in which anyone can apply to buy the bonds, or a *private offer*, where the customers of the investment bank are offered the stock. The secondary market is the market in which existing bonds are subsequently traded.

Bond markets are regulated as part of the overall financial system. In most countries there is an independent regulator responsible for overseeing both the structure of the market and the bona fides of market participants. For instance, the US market regulator is the Securities and Exchange Commission (SEC). The UK regulator, the Financial Services Authority (FSA), is responsible for regulating both wholesale and retail markets; for example, it reviews the capital requirements for commercial and investment banks, and it is also responsible for regulating the retail mortgage market. Money markets are usually overseen by the country's central bank – for example, the Federal Reserve manages the daily money supply in the United States, while the Bank of England provides liquidity to the market by means of its daily money market repo operation.

The government bond market

Government bonds are traded on the following four markets: in addition to the primary and secondary markets, we have the *when-issued market* and the *repo market*.

- The primary market: newly issued securities are first sold through an auction, which is conducted on a competitive bid basis. The auction process happens between the government and primary/non-primary dealers according to regular cycles for securities with specific maturities.[3]
- The secondary market: here a group of government securities dealers offer continuous bid and ask prices on specific outstanding government bonds. This is an OTC market.
- The when-issued market: here securities are traded on a forward basis before they are issued by the government.
- The repo market: in this market securities are used as collateral for loans. A distinction must be made between the *general-collateral* repo rate (GC) and the *special* repo rate. The GC repo rate applies to the major part of government securities. Special repo rates are specific repo rates. They typically concern on-the-run and cheapest-to-deliver securities, which are very expensive. This is the reason that special repo rates are at a level below the GC repo rate. Indeed, as these securities are very much in demand, the borrowers of these securities on the repo market receive a relatively lower repo rate compared to normal government securities.

The bonds issued by regional governments and certain public sector bodies, such as national power and telecommunications utilities, are usually included as 'government' debt, as they are almost always covered by an explicit or implicit government guarantee. All other categories of borrower are therefore deemed to be 'corporate' borrowers. Generally the term 'corporate markets' is used to cover bonds issued by non-government borrowers.

The corporate bond market

In the context of a historically low level of interest rates, linked to a decreasing trend in inflation as well as in budget deficits, the corporate bond market is rapidly developing and growing. This strong tendency affects both the supply and the demand. While corporate supply is expanding, in relation to bank disintermediation, corporate demand is rising as more and more investors accustomed to dealing with only government bonds are including corporate bonds in their portfolios so as to capture spread and generate performance.

The market by country and sector

Within the four major bond markets in the world, the US dollar (USD) corporate market is the most mature, followed by the sterling (GBP) market and the euro

[3]The US auction cycles are as follows. Two-year notes are auctioned every month and settled on the 15th. Five-year notes are auctioned quarterly, in February, May, August and November of each year, and settled at the end of the month. Ten-year notes are auctioned quarterly, in February, May, August and November of each year, and settled on the 15th of the month. Thirty-year bonds are auctioned semi-annually, in February and August of each year, and settled on the 15th of the month. Auctions are announced by the Treasury one week in advance, the issuing date being set one to five days after the auction.

(EUR) market, the growth of the latter being reinforced by the launching of the euro. The Japanese yen (JPY) market differentiates itself from the others, because of the credit crunch situation and economic difficulties it has been facing since the Asian crisis. The USD corporate bond market is the largest and most diversified: it is for instance more than twice as big as the Euro market, and low investment-grade ratings are much more represented (being over 80% of the index).

The corporate bond market can be divided into three main sectors: financial, industrial and utility. Apart from the USD market, the financial sector is over-represented. It is another proof of the maturity of the USD market, where the industrial sector massively uses the market channel in order to finance investment projects. It is also worth noting that the sector composition in the USD market is far more homogeneous than in the other markets. For example, the banking sector is systematically predominant in the GBP, EUR and JPY financial markets, while the telecommunication sector exceeds one-third of the Euro industrial market. As a result, local credit portfolio diversification can be better achieved in the USD market than in the others.

Underwriting a new issue

The issue of corporate debt in the capital markets requires a primary market mechanism. The first requirement is a collection of merchant banks or investment banks that possess the necessary expertise. Investment banks provide advisory services on corporate finance as well as underwriting services, which is a guarantee to place an entire bond issue into the market in return for a fee. As part of the underwriting process the investment bank will either guarantee a minimum price for the bonds, or aim to place the paper at the best price available. The major underwriting institutions in emerging economies are often branch offices of the major integrated global investment banks.

Small-size bond issues may be underwritten by a single bank. It is common, however, for larger issues, or issues that are aimed at a cross-border investor base, to be underwritten by a syndicate of investment banks. This is a group of banks that collectively underwrite a bond issue, with each syndicate member being responsible for placing a proportion of the issue. The bank that originally won the mandate to place the paper invites other banks to join the syndicate. This bank is known as the *lead underwriter, lead manager* or *book-runner*. An issue is brought to the market simultaneously by all syndicate members, usually via the *fixed price re-offer* mechanism.[4] This is designed to guard against some

[4]In a *fixed price re-offer* scheme the lead manager will form the syndicate, which will agree on a fixed issue price, a fixed commission and the distribution amongst themselves of the quantity of bonds they will take as part of the syndicate. The banks then re-offer the bonds that they have been allotted to the market, at the agreed price. This technique gives the lead manager greater control over an issue. It sets the price at which other underwriters in the syndicate can initially sell the bonds to investors. The fixed price re-offer mechanism is designed to prevent underwriters from selling the bonds back to the lead manager at a discount to the original issue price, that is, 'dumping' the bonds.

syndicate members in an offering selling stock at a discount in the *grey market*, to attract investors.[5] This would force the lead manager to buy the bonds back if it wished to support the price. Under the fixed price re-offer method, price undercutting is not possible as all banks are obliged not to sell their bonds below the initial offer price that has been set for the issue. The fixed price usually is in place up to the first settlement date, after which the bond is free to trade in the secondary market.

The Eurobond market

The key feature of a Eurobond is the way it is issued, internationally across borders and by an international underwriting syndicate. The method of issuing Eurobonds reflects the cross-border nature of the transaction, and unlike government markets where the auction is the primary issue method, Eurobonds are typically issued under a fixed price re-offer method or a *bought deal*.[6]

The range of borrowers in the Euromarkets is very diverse. From virtually the inception of the market, borrowers representing corporates, sovereign and local governments, nationalised corporations, supranational institutions and financial institutions have raised finance in the international markets. The majority of borrowing has been by national governments, regional governments and public agencies of developed countries, although the Eurobond market is increasingly a source of finance for developing country governments and corporates.

Governments and institutions access the Euromarkets for a number of reasons. Under certain circumstances it is more advantageous for a borrower to raise funds outside its domestic market, due to the effects of tax or regulatory rules.[7] The international markets are very competitive in terms of using intermediaries, so a borrower may well be able to raise cheaper funds in the international markets.

---

[5]The *grey market* is a term used to describe trading in the bonds before they officially come to the market, mainly market makers selling the bond short to other market players or investors. Activity in the grey market serves as useful market intelligence to the lead manager, who can gauge the level of demand that exists in the market for the issue. A final decision on the offer price is of course made on the issue day itself.

[6]In a *bought deal*, a lead manager or a managing group approaches the issuer with a firm bid, specifying issue price, amount, coupon and yield. Only a few hours are allowed for the borrower to accept or reject the terms. If the bid is accepted, the lead manager purchases the entire bond issue from the borrower. The lead manager then has the option of selling part of the issue to other banks for distribution to investors, or doing so itself. In a volatile market the lead manager will probably parcel some of the issue to other banks for placement. However, it is at this time that the risk of banks dumping bonds on the secondary market is highest; in this respect lead managers will usually pre-place the bonds with institutional investors before the bid is made. The bought deal is focused primarily on institutional rather than private investors. As the syndicate process is not used, the bought deal requires a lead manager with sufficient capital and placement power to enable the entire issue to be placed.

[7]There is no formal regulation of the Eurobond market as such, but each market participant will be subject to the regulation of its country regulator.

Other reasons why borrowers access Eurobond markets include:

- A desire to diversify sources of long-term funding. A bond issue is often placed with a wide range of institutional and private investors, rather than the more restricted investor base that may prevail in a domestic market. This gives the borrower access to a wider range of lenders, and for corporate borrowers this also enhances the international profile of the company.
- For both corporates and emerging country governments, the prestige associated with an issue of bonds in the international market.
- The flexibility of a Eurobond issue compared with a domestic bond issue or bank loan, illustrated by the different types of Eurobond instruments available.

Against this are balanced the potential downsides of a Eurobond issue, which include the following:

- For all but the largest and most creditworthy of borrowers, the rigid nature of the issue procedure becomes significant during times of interest-rate and exchange-rate volatility, reducing the funds available for borrowers.
- Issuing debt in currencies other than those in which a company holds matching assets, or in which there are no prospects of earnings, exposes the issuer to foreign exchange risk.

Table 4.3 shows some issues in the Eurobond market during 2009. The market remains an efficient and attractive market in which a company can raise finance for a wide range of maturities. The institutional investors include insurance companies, pension funds, investment trusts, commercial banks, and corporations – just as in domestic corporate bond markets. Other investors include central banks and government agencies; for example, the Kuwait Investment Office and the Saudi Arabian Monetary Agency both have large Eurobond holdings. In the UK, banks and securities houses are keen holders of Eurobonds issued by other financial institutions.

Credit risk

As is the case for government and municipal bonds, the issuer of a corporate bond has the obligation to honour its commitments to the bondholder. A failure to pay back interests or principal according to the terms of the agreement constitutes what is known as *default*. Basically, there are two sources of default. First, the shareholders of a corporation can decide to break the debt contract. This comes from their limited liability status: they are liable for the corporation's losses only up to their investment in it. They do not have to pay back their creditors when it affects their personal wealth. Second, creditors can prompt bankruptcy when specific debt protective clauses, known as covenants, are infringed.

In case of default, there are typically three eventualities:

- First, default can lead to immediate bankruptcy. Depending on the seniority and face value of their debt securities, creditors are fully, partially or not paid

Table 4.3 Selected euro-denominated Eurobond issues in 2009

ISSUER	*CPN %*	*S&P*	*MOODYS*	*MATURITY*	*ISSUED AMOUNT EUR*	*LAUNCH SPREAD (BPS)*
ERSTE GROUP BANK AG	3	AAA	Aaa	23/01/2012	1,500,000,000.00	147.9
DANSKE BANK A/S	2.5	AAA	Aaa	21/09/2010	3,000,000,000.00	109.4
BANK OF SCOTLAND PLC	3.875	AAA	Aaa	05/11/2010	3,000,000,000.00	136.3
LAFARGE SA	7.625	BBB−	Baa3	27/05/2014	1,000,000,000.00	532.6
POHJOLA BANK PLC	4.5	AA−	Aa1	22/05/2014	750,000,000.00	212.5
SWISS RE TREASURY (US)	7	A+	A1	19/05/2014	600,000,000.00	449.9
IRISH NATIONWIDE BLDG SO	3.5	AA+	Aaa	22/09/2010	750,000,000.00	247.4
DEUTSCHE TELEKOM INT FIN	4.375	BBB+	Baa1	02/06/2014	500,000,000.00	181.6
DNB NOR BANK ASA	4.5	AA−	Aa1	29/05/2014	2,000,000,000.00	201.8
VOLVO TREASURY AB	7.875	BBB+	Baa1	01/10/2012	500,000,000.00	592.5
COMPAGNIE DE ST GOBAIN	6	BBB+	Baa1	20/05/2013	750,000,000.00	375.3

Source: Bloomberg L.P.

back thanks to the sale of the firm's assets. The percentage of the interests and principal they receive, according to seniority, is called the *recovery rate*.

- Second, default can result in a reorganisation of the firm within a formal legal framework. For example, under Chapter 11 of the American law, corporations that are in default are granted a deadline so as to overcome their financial difficulties. This depends on the country's legislation.
- Third, default can lead to an informal negotiation between shareholders and creditors. This results in an exchange offer through which shareholders propose to creditors the exchange of their old debt securities for a package of cash and newly issued securities.

A corporate debt issue is priced over the same currency government bond yield curve. A liquid benchmark yield curve therefore is required to facilitate pricing. The extent of a corporate bond's yield spread over the government yield curve is a function of the market's view of the credit risk of the issuer (for which formal credit ratings are usually used) and the perception of the liquidity of the issue. The pricing of corporate bonds is sometimes expressed as a spread over the equivalent maturity government bond, rather than as an explicit stated yield, or sometimes as a spread over another market reference index such as Libor. Figure 4.2 illustrates some typical yield spreads for different ratings and maturities of corporate bonds.

Corporate bonds are much affected by credit risk. Their yields normally contain a default premium over government bonds, accounting for total default or credit risk, as well as over swaps. Swap spread, that is, the difference between the swap yield and the government yield with same maturity, is regarded as a systematic credit premium. In the four main bond markets swap yields reflect bank risk with rating AA, that is, the first rating grade below AAA, the normal rating for government bonds, accounting for specific default or credit risk.

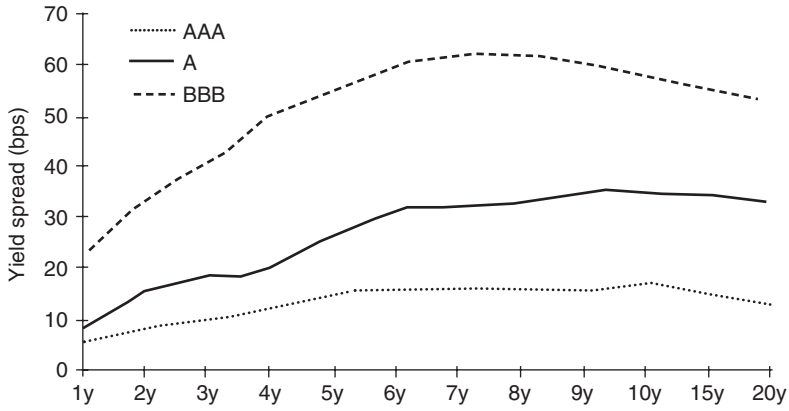


Figure 4.2 Yield spread by rating and maturity

Formal credit ratings are important in the corporate markets. Investors usually use a domestic rating agency in conjunction with an established international agency such as Moody's or Standard & Poor's. As formal ratings are viewed as important by investors, it is in the interest of issuing companies to seek a rating from an established agency, especially if it is seeking to issue foreign currency and/or place its debt across national boundaries. Generally Eurobond issuers are investment-grade rated, and only a very small number are not rated at all.

Treasury securities are considered to have no credit risk. The interest rates they bear are the key interest rates in the United States as well as in international capital markets. Agency securities' debt is high-quality debt. As a matter of fact, all rated agency senior debt issues are triple-A rated by Moody's and Standard & Poor's. This rating most often reflects healthy financial fundamentals and sound management, but also and above all the agencies' relationship to the US government. Among the numerous legal characteristics of the government agencies' debt, one can find that:

- agencies' directors are appointed by the President of the United States
- issuance is only upon approval by the US Treasury
- securities are issuable and payable through the Federal Reserve System
- securities are eligible collateral for Federal Reserve Bank advances and discounts
- securities are eligible for open market purchases.

Municipal debt issues, when rated, carry ratings ranging from triple-A, for the best ones, to C or D, for the worst ones. Four basic criteria are used by rating agencies to assess municipal bond ratings:

- the issuer's debt structure
- the issuer's ability and political discipline for maintaining sound budgetary operations
- the local tax and intergovernmental revenue sources of the issuer
- the issuer's overall socio-economic environment.

Pricing and yield

Bonds are debt capital market instruments that represent a cash flow payable during a specified time period heading into the future. This cash flow represents the interest payable on the loan and the loan redemption. So a bond essentially is a loan, albeit one that is tradable in a secondary market. This differentiates bond market securities from commercial bank loans.

For some considerable time the analysis of bonds was frequently presented in what might be termed 'traditional' terms, with description limited to gross redemption yield or *yield to maturity*. However these days basic bond maths analysis is presented in slightly different terms, as described in a range of books and

articles such as those by Ingersoll (1987), Shiller (1990), Neftci (1996), Jarrow (1996), Van Deventer (1997) and Sundaresan (1997), among others.[8] For this reason we review the basic elements in this chapter but then consider the academic approach and description of bond pricing, and a review of the term structure of interest rates. Interested readers may wish to consult the texts in the bibliography for further information.

In the analysis that follows bonds are assumed to be *default-free*, which means that there is no possibility that the interest payments and principal repayment will not be made. Such an assumption is accurate when one is referring to government bonds such as US Treasuries, UK gilts and so on. However it is unreasonable when applied to bonds issued by corporates or lower-rated sovereign borrowers. Nevertheless it is still relevant to understand the valuation and analysis of bonds that are default-free, as the pricing of bonds that carry default risk is based on the price of risk-free government securities. Essentially the price investors charge borrowers that are not of risk-free credit standing is the price of government securities plus some *credit risk* premium.

Bond pricing and yield: the traditional approach

Bond pricing

The interest rate that is used to discount a bond's cash flows (therefore called the *discount* rate) is the rate required by the bondholder. It is therefore known as the bond's *yield*. The yield on the bond will be determined by the market, and is the price demanded by investors for buying it, which is why it is sometimes called the bond's *return*. The required yield for any bond will depend on a number of political and economic factors, including what yield is being earned by other bonds of the same class. Yield is always quoted as an annualised interest rate, so that for a semi-annually coupon paying bond exactly half of the annual rate is used to discount the cash flows.

The *fair price* of a bond is the present value of all its cash flows. Therefore when pricing a bond we need to calculate the present value of all the coupon interest payments and the present value of the redemption payment, and sum these. The price of a conventional bond that pays annual coupons can therefore be given by (4.1).

$$\begin{aligned} P &= \frac{C}{(1+r)} + \frac{C}{(1+r)^2} + \frac{C}{(1+r)^3} + \cdots \frac{C}{(1+r)^N} + \frac{M}{(1+r)^N} \\ &= \sum_{n=1}^{N} \frac{C}{(1+r)^n} + \frac{M}{(1+r)^N} \end{aligned} \tag{4.1}$$

[8] These and other recommended readings are given at the end of this chapter.

where

P is the price
C is the annual coupon payment (for semi-annual coupons, will be $C/2$)
r is the discount rate (therefore, the required yield)
N is the number of years to maturity (therefore, the number of interest periods in an annually-paying bond; for a semi-annual bond the number of interest periods is $N \times 2$).
M is the maturity payment or par value (usually 100% of currency)

For long-hand calculation purposes the first half of (4.1) is usually simplified and is sometimes encountered in one of the two ways shown in (4.2).

$$P = C\left[\frac{1-\left[\frac{1}{(1-r)^N}\right]}{r}\right] \tag{4.2}$$

or

$$P = \frac{C}{r}\left[1-\left[\frac{1}{(1+r)^N}\right]\right]$$

The price of a bond that pays semi-annual coupons is given by the expression at (4.3), which is our earlier expression modified to allow for the twice-yearly discounting:

$$\begin{aligned} P &= \frac{C/2}{(1+\frac{1}{2}r)} + \frac{C/2}{(1+\frac{1}{2}r)^2} + \frac{C/2}{(1+\frac{1}{2}r)^3} + \cdots \frac{C/2}{(1+\frac{1}{2}r)^{2N}} + \frac{M}{(1+\frac{1}{2}r)^{2N}} \\ &= \sum_{n=1}^{2N} \frac{C/2}{(1+\frac{1}{2}r)^n} + \frac{M}{(1+\frac{1}{2}r)^{2N}} \\ &= \frac{C}{r}\left[1-\frac{1}{(1+\frac{1}{2}r)^{2N}}\right] + \frac{M}{(1+\frac{1}{2}r)^{2N}} \end{aligned} \tag{4.3}$$

Note how we set $2N$ as the power to which to raise the discount factor, as there are two interest payments every year for a bond that pays semi-annually. Therefore a more convenient function to use might be the number of interest periods in the life of the bond, as opposed to the number of years to maturity, which we could set as n, allowing us to alter the equation for a semi-annually paying bond as:

$$P = \frac{C}{r}\left[1-\frac{1}{(1+\frac{1}{2}r)^n}\right] + \frac{M}{(1+\frac{1}{2}r)^n} \tag{4.4}$$

The formula at (4.4) calculates the fair price on a coupon payment date, so that there is no *accrued interest* incorporated into the price. It also assumes that there is an even number of coupon payment dates remaining before maturity. The concept of accrued interest is an accounting convention, and treats coupon interest as accruing every day that the bond is held. This amount is added to the discounted present value of the bond (the *clean price*) to obtain the market value of the bond, known as the *dirty price*.

The date used as the point for calculation is the settlement date for the bond, the date on which a bond will change hands after it is traded. For a new issue of bonds the settlement date is the day when the stock is delivered to investors and payment is received by the bond issuer. The settlement date for a bond traded in the *secondary market* is the day that the buyer transfers payment to the seller of the bond and the seller transfers the bond to the buyer. Different markets will have different settlement conventions, for example UK gilts normally settle one business day after the trade date (the notation used in bond markets is 'T + 1') whereas Eurobonds settle on T + 3. The term *value date* is sometimes used in place of settlement date. The two terms are, however, not strictly synonymous. A settlement date can only fall on a business date, so that a gilt traded on a Friday will settle on a Monday. A value date can however sometimes fall on a non-business day, for example when accrued interest is being calculated.

If there is an odd number of coupon payment dates before maturity the formula at (4.4) is modified as shown in (4.5).

$$P = \frac{C}{r}\left[1 - \frac{1}{(1+\frac{1}{2}r)^{2N+1}}\right] + \frac{M}{(1+\frac{1}{2}r)^{2N+1}} \tag{4.5}$$

The standard formula also assumes that the bond is traded for settlement on a day that is precisely one interest period before the next coupon payment. The price formula is adjusted if dealing takes place in-between coupon dates. If we take the *value date* for any transaction, we then need to calculate the number of calendar days from this day to the next coupon date. We then use the following ratio i when adjusting the exponent for the discount factor:

$$i = \frac{\text{Days from value date to next coupon date}}{\text{Days in the interest payment}}$$

The number of days in the interest period is the number of calendar days between the last coupon date and the next one, and it will depend on the day count basis used for that specific bond. The price formula is then modified as shown at (4.6).

$$P = \frac{C}{(1+r)^{i}} + \frac{C}{(1+r)^{1+i}} + \frac{C}{(1+r)^{2+i}} + \cdots + \frac{C}{(1+r)^{n-1+i}} + \frac{M}{(1+r)^{n-1+i}} \tag{4.6}$$

where the variables C, M, n and r are as before. Note that (4.6) assumes r for an annually-paying bond and is adjusted to $r/2$ for a semi-annual coupon paying bond.

There also exist *perpetual* or *irredeemable* bonds which have no redemption date, so that interest on them is paid indefinitely. They are also known as *undated* bonds. An example of an undated bond is the 3½% War Loan, a UK gilt originally issued in 1916 to help pay for the 1914–18 war effort. Most undated bonds date from a long time in the past and it is unusual to see them issued today. In structure the cash flow from an undated bond can be viewed as a continuous annuity. The fair price of such a bond is given from 2.11 by setting $N = \infty$, such that:

$$\mathbf{P} = \frac{C}{r} \tag{4.7}$$

In most markets bond prices are quoted in decimals, in minimum increments of 1/100ths. This is the case with Eurobonds, euro-denominated bonds and gilts, for example. Certain markets, the US Treasury market, South African and Indian government bonds for example, quote prices in *ticks*, where the minimum increment is 1/32nd. One tick is therefore equal to 0.03125. A US Treasury might be priced at '98-05' which means '98 and 5 ticks'. This is equal to 98 and 5/32nds which is 98.15625.

Bonds that do not pay a coupon during their life are known as *zero-coupon* bonds or *strips*, and the price for these bonds is determined by modifying (4.1) to allow for the fact that $C = 0$. We know that the only cash flow is the maturity payment, so we may set the price as:

$$P = \frac{M}{(1+r)^N} \tag{4.8}$$

where M and r are as before and N is the number of years to maturity. The important factor is to allow for the same number of interest periods as coupon bonds of the same currency. That is, even though there are no actual coupons, we calculate prices and yields on the basis of a *quasi-coupon* period. For a US dollar or a sterling zero-coupon bond, a five-year zero-coupon bond would be assumed to cover ten quasi-coupon periods, which would set the price equation as:

$$P = \frac{M}{(1+\frac{1}{2}r)^{2N}} \tag{4.9}$$

We have to note carefully the quasi-coupon periods in order to maintain consistency with conventional bond pricing.

An examination of the bond price formula tells us that the yield and price for a bond are closely related. A key aspect of this relationship is that the price changes in the opposite direction to the yield. This is because the price of the bond is the net present value of its cash flows; if the discount rate used in the present value calculation increases, the present values of the cash flows will decrease. This occurs whenever the yield level required by bondholders increases. In the same way if the required yield decreases, the price of the bond will rise.

Bond yield

We have observed how to calculate the price of a bond using an appropriate discount rate known as the bond's *yield*. We can reverse this procedure to find the yield of a bond where the price is known, which would be equivalent to calculating the bond's *internal rate of return* (IRR). The IRR calculation is taken to be a bond's *yield to maturity* or *redemption yield*, and is one of various yield measures used in the markets to estimate the return generated from holding a bond. In most markets bonds are generally traded on the basis of their prices, but because of the complicated patterns of cash flows that different bonds can have, they are generally compared in terms of their yields. This means that a market-maker will usually quote a two-way price at which she will buy or sell a particular bond, but it is the yield at which the bond is trading that is important to the market-maker's customer. This is because a bond's price does not actually tell us anything useful about what we are getting. Remember that in any market there will be a number of bonds with different issuers, coupons and terms to maturity. Even in a homogenous market such as the gilt market, different gilts will trade according to their own specific characteristics. To compare bonds in the market therefore we need the yield on any bond and it is yields that we compare, not prices.

The yield on any investment is the interest rate that will make the present value of the cash flows from the investment equal to the initial cost (price) of the investment. Mathematically the yield on any investment, represented by r, is the interest rate that satisfies equation (4.10) below, which is simply the bond price equation we have already reviewed.

$$P = \sum_{n=1}^{N} \frac{C}{(1+r)^n} \tag{4.10}$$

But as we have noted there are other types of yield measure used in the market for different purposes. The simplest measure of the yield on a bond is the *current yield,* also know as the *flat yield, interest yield* or *running yield*. The running yield is given by (4.11).

$$rc = \frac{C}{P} \times 100 \tag{4.11}$$

where rc is the current yield.

In (4.11) C is not expressed as a decimal. Current yield ignores any capital gain or loss that might arise from holding and trading a bond, and does not consider the time value of money. It essentially calculates the bond coupon income as a proportion of the price paid for the bond, and to be accurate would have to assume that the bond was more like an annuity than a fixed-term instrument.

The current yield is useful as a 'rough-and-ready' interest rate calculation; it is often used to estimate the cost of or profit from a short-term holding of a bond. For example if other short-term interest rates such as the one-week or three-month rates are higher than the current yield, holding the bond is said to involve a *running cost*. This is also known as *negative carry* or *negative funding*. The term is

used by bond traders and market makers and *leveraged* investors. The *carry* on a bond is a useful measure for all market practitioners as it illustrates the cost of holding or *funding* a bond. The funding rate is the bondholder's short-term cost of funds. A private investor could also apply this to a short-term holding of bonds.

The *yield to maturity* (YTM) or *gross redemption yield* is the most frequently used measure of return from holding a bond.[9] Yield to maturity takes into account the pattern of coupon payments, the bond's term to maturity and the capital gain (or loss) arising over the remaining life of the bond. We saw from our bond price formula on page 68 that these elements were all related, and were important components determining a bond's price. If we set the IRR for a set of cash flows to be the rate that applies from a start date to an end date we can assume the IRR to be the YTM for those cash flows. The YTM therefore is equivalent to the *internal rate of return* on the bond, the rate that equates the value of the discounted cash flows on the bond to its current price. The calculation assumes that the bond is held until maturity and therefore it is the cash flows to maturity that are discounted in the calculation. It also employs the concept of the time value of money.

As we would expect, the formula for YTM is essentially that for calculating the price of a bond. For a bond paying annual coupons the YTM is calculated by solving equation (4.1). Note that the expression at (4.1) has two variable parameters, the price P and yield r. It cannot be rearranged to solve for yield r explicitly, and in fact the only way to solve for the yield is to use the process of numerical iteration. The process involves estimating a value for r and calculating the price associated with the estimated yield. If the calculated price is higher than the price of the bond at the time, the yield estimate is lower than the actual yield, and so it must be adjusted until it converges to the level that corresponds with the bond price.[10] For the YTM of a semi-annual coupon bond we have to adjust the formula to allow for the semi-annual payments, shown at (4.3).

To differentiate redemption yield from other yield and interest rate measures described in this book, we henceforth refer to it as *rm*.

Example 4.1: Yield to maturity for semi-annual coupon bond

A semi-annual paying bond has a price of £98.50, an annual coupon of 6% and there is exactly one year before maturity. The bond therefore has three remaining cash flows, comprising two coupon payments of £3 each and a redemption payment of £100. Equation (4.10) can be used with the following inputs:

$$98.50 = \frac{3.00}{(1+\frac{1}{2}rm)} + \frac{103.00}{(1-\frac{1}{2}rm)^2}$$

[9]In this book the terms *yield to maturity* and *gross redemption yield* are used synonymously. The latter term is encountered in sterling markets.

[10]Bloomberg also uses the term *yield-to-workout* where 'workout' refers to the maturity date for the bond.

Note that we use half of the YTM value *rm* because this is a semi-annual paying bond. The expression above is a quadratic equation, which is solved using the standard solution for quadratic equations, which is noted below.

$$ax^2 + bx + c = 0$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

In our expression if we let $x = (1 + rm/2)$, we can rearrange the expression as follows:

$$98.50x^2 - 3.0x - 103.00 = 0$$

We then solve for a standard quadratic equation, and as such there will be two solutions, only one of which gives a positive redemption yield. The positive solution is $rm/2 = 0.037929$ so that $rm = 7.5859\%$.

As an example of the iterative solution method, suppose that we start with a trial value for *rm* of $r_1 = 7\%$ and plug this into the right-hand side of equation (4.10). This gives a value for the right-hand side of:

$$\text{RHS}_1 = 99.050$$

which is higher than the left-hand side (LHS = 98.50); the trial value for *rm* was therefore too low. Suppose then that we try next $r_2 = 8\%$ and use this as the right-hand side of the equation. This gives:

$$\text{RHS}_2 = 98.114$$

which is lower than the LHS. Because RHS_1 and RHS_2 lie on either side of the LHS value we know that the correct value for *rm* lies between 7% and 8%. Using the formula for linear interpolation,

$$rm = r_1 + (r_2 - r_1)\frac{RHS_1 - LHS}{RHS_1 - RHS_2}$$

our linear approximation for the redemption yield is $rm = 7.587\%$, which is near the exact solution.

Note that the redemption yield as calculated as discussed in this section is the *gross redemption yield*, the yield that results from payment of coupons without deduction of any withholding tax. The *net redemption yield* is obtained by multiplying the coupon rate *C* by (1 – marginal tax rate). The net yield is what will be received if the bond is traded in a market where bonds pay coupon *net*, which

means net of a withholding tax. The net redemption yield is always lower than the gross redemption yield.

We have already alluded to the key assumption behind the YTM calculation, namely that the rate *rm* remains stable for the entire period of the life of the bond. By assuming the same yield we can say that all coupons are reinvested at the same yield *rm*. For the bond in Example 4.1 this means that if all the cash flows are discounted at 7.59% they will have a total net present value of 98.50. This is patently unrealistic since we can predict with virtual certainty that interest rates for instruments of similar maturity to the bond at each coupon date will not remain at this rate for the life of the bond. In practice however, investors require a rate of return that is equivalent to the price that they are paying for a bond and the redemption yield is, to put it simply, as good a measurement as any. A more accurate measurement might be to calculate present values of future cash flows using the discount rate that is equal to the market's view on where interest rates will be at that point, known as the *forward* interest rate. However forward rates are *implied* interest rates, and a YTM measurement calculated using forward rates can be as speculative as one calculated using the conventional formula. This is because the *actual* market interest rate at any time is invariably different from the rate implied earlier in the forward markets. So a YTM calculation made using forward rates would not be realised in practice either.[11] We shall see later how the *zero-coupon* interest rate is the true interest rate for any term to maturity. However the YTM is, despite the limitations presented by its assumptions, the main measure of return used in the markets.

We have noted the difference between calculating redemption yield on the basis of both annual and semi-annual coupon bonds. Analysis of bonds that pay semi-annual coupons incorporates semi-annual discounting of semi-annual coupon payments. This is appropriate for most UK and US bonds. However government bonds in most of continental Europe and most Eurobonds pay annual coupon payments, and the appropriate method of calculating the redemption yield is to use annual discounting. The two yields measures are not therefore directly comparable. We could make a Eurobond directly comparable with a UK gilt by using semi-annual discounting of the Eurobond's annual coupon payments. Alternatively we could make the gilt comparable with the Eurobond by using annual discounting of its semi-annual coupon payments. The price/yield formulae for different discounting possibilities we encounter in the markets are listed below. (As usual we assume that the calculation takes place on a coupon payment date so that accrued interest is zero.)

Semi-annual discounting of annual payments:

$$P_d = \frac{C}{(1+\frac{1}{2}rm)^2} + \frac{C}{(1+\frac{1}{2}rm)^4} + \frac{C}{(1+\frac{1}{2}rm)^6} + \cdots + \frac{C}{(1+\frac{1}{2}rm)^{2N}} + \frac{M}{(1+\frac{1}{2}rm)^{2N}} \quad (4.12)$$

[11] Such an approach is used to price interest-rate swaps, however.

Annual discounting of semi-annual payments:

$$P_d = \frac{C/2}{(1+rm)^{1/2}} + \frac{C/2}{(1+rm)} + \frac{C/2}{(1+rm)^{3/2}} + \cdots + \frac{C/2}{(1+rm)^N} + \frac{M}{(1+rm)^N} \tag{4.13}$$

Consider a bond with a dirty price of 97.89, a coupon of 6% and five years to maturity. This bond would have the following gross redemption yields under the different yield calculation conventions:

Discounting	*Payments*	*Yield to maturity (%)*
Semi-annual	Semi-annual	6.500
Annual	Annual	6.508
Semi-annual	Annual	6.428
Annual	Semi-annual	6.605

This proves what we have already observed, namely that the coupon and discounting frequency will impact the redemption yield calculation for a bond. We can see that increasing the frequency of discounting will lower the yield, while increasing the frequency of payments will raise the yield. When comparing yields for bonds that trade in markets with different conventions it is important to convert all the yields to the same calculation basis.

Intuitively we might think that doubling a semi-annual yield figure will give us the annualised equivalent. In fact this will result in an inaccurate figure due to the multiplicative effects of discounting, and one that is an underestimate of the true annualised yield. The correct procedure for producing an annualised yield from semi-annual and quarterly yields is given by the expressions below.

The general conversion expression is given by (4.14):

$$rm_a = (1 + \text{interest rate})^m - 1 \tag{4.14}$$

where m is the number of coupon payments per year.

Specifically we can convert between yields using the expressions given at (4.15) and (4.16):

$$rm_a = [(1 + \tfrac{1}{2}rm_s)^2 - 1] \tag{4.15}$$

$$rm_s = [(1 + rm_a)^{1/2} - 1] \times 2$$

$$rm_a = [(1 + \tfrac{1}{4}rm_q)^4 - 1] \tag{4.16}$$

$$rm_q = [(1 + rm_a)^{1/4} - 1] \times 4$$

where rm_q, rm_s and rm_a are respectively the quarterly, semi-annually and annually compounded yields to maturity.

Example 4.2

A UK gilt paying semi-annual coupons and a maturity of 10 years has a quoted yield of 4.89%. A European government bond of similar maturity is quoted at a yield of 4.96%. Which bond has the higher effective yield?

The effective annual yield of the gilt is:

$$rm = (1 + \text{½}0..0489)^2 - 1 = 4.9498\%$$

Therefore the gilt does indeed have the lower yield.

The market convention is sometimes simply to double the semi-annual yield to obtain the annualised yields, despite the fact that this produces an inaccurate result. It is only acceptable to do this for rough calculations. An annualised yield obtained by multiplying the semi-annual yield by two is known as a *bond equivalent yield*.

The major disadvantage of the YTM measure has already been alluded to. Another disadvantage of the yield to maturity measure of return is where investors do not hold bonds to maturity. The redemption yield will not be great where the bond is not being held to redemption. Investors might then be interested in other measures of return, which we can look at later.

To reiterate, the redemption yield measure assumes that:

- the bond is held to maturity
- all coupons during the bond's life are reinvested at the same (redemption yield) rate.

Therefore the YTM can be viewed as an *expected* or *anticipated* yield, and is closest to reality perhaps where an investor buys a bond on first issue and holds it to maturity. Even then the actual realised yield on maturity would be different from the YTM figure because of the inapplicability of the second condition above.

In addition, as coupons are discounted at the yield specific for each bond, it actually becomes inaccurate to compare bonds using this yield measure. For instance the coupon cash flows that occur in two years' time from both a two-year and five-year bond will be discounted at different rates (assuming we do not have a flat yield curve). This would occur because the YTM for a five-year bond is invariably different to the YTM for a two-year bond. However it would clearly not be correct to discount a two-year cash flow at different rates, because we can see that the present value calculated today of a cash flow in two years' time should be the same whether it is sourced from a short or long-dated bond. Even if the first condition noted above for the YTM calculation is satisfied, it is clearly unlikely for any but the shortest maturity bond that all coupons will be reinvested at the same rate. Market interest rates are in a state of constant flux and hence this would

affect money reinvestment rates. Therefore although yield to maturity is the main market measure of bond levels, it is not a true interest rate. This is an important result and we shall explore the concept of a true interest rate in Chapter 7.

Accrued interest, clean and dirty bond prices

The consideration of bond pricing up to now has ignored coupon interest. All bonds (except zero-coupon bonds) accrue interest on a daily basis, and this is then paid out on the coupon date. The calculation of bond prices using present value analysis does not account for coupon interest or *accrued interest*. In all major bond markets the convention is to quote price as a clean price. This is the price of the bond as given by the net present value of its cash flows, but excluding coupon interest that has accrued on the bond since the last dividend payment. As all bonds accrue interest on a daily basis, even if a bond is held for only one day, interest will have been earned by the bondholder. However we have referred already to a bond's *all-in* price, which is the price that is actually paid for the bond in the market. This is also known as the *dirty price* (or *gross price*), which is the clean price of a bond plus accrued interest. In other words the accrued interest must be added to the quoted price to get the total consideration for the bond.

Accruing interest compensates the seller of the bond for giving up all of the next coupon payment even though she will have held the bond for part of the period since the last coupon payment. The clean price for a bond will move with changes in market interest rates; assuming that this is constant in a coupon period, the clean price will be constant for this period. However the dirty price for the same bond will increase steadily from one interest payment date to the next one. On the coupon date the clean and dirty prices are the same and the accrued interest is zero. Between the coupon payment date and the next *ex dividend* date the bond is traded *cum dividend*, so that the buyer gets the next coupon payment. The seller is compensated for not receiving the next coupon payment by receiving accrued interest instead. This is positive and increases up to the next ex dividend date, at which point the dirty price falls by the present value of the amount of the coupon payment. The dirty price at this point is below the clean price, reflecting the fact that accrued interest is now negative. This is because after the ex dividend date the bond is traded 'ex dividend'; the seller not the buyer receives the next coupon and the buyer has to be compensated for not receiving the next coupon by means of a lower price for holding the bond.

The net interest accrued since the last ex dividend date is determined as follows:

$$AI = C \times \left[\frac{N_{xt} - N_{xc}}{DayBase} \right] \tag{4.17}$$

where

AI is the net accrued interest
C is the bond coupon

N_{xc} is the number of days between the ex dividend date and the coupon payment date (seven business days for UK gilts)

N_{xt} is the number of days between the ex dividend date and the date for the calculation

DayBase is the day count base (365 or 360)

Certain bonds do not have an ex-dividend period, for example Eurobonds, and accrue interest right up to the coupon date.

Interest accrues on a bond from and including the last coupon date up to and excluding what is called the *value date*. The value date is almost always the *settlement date* for the bond, or the date when a bond is passed to the buyer and the seller receives payment. Interest does not accrue on bonds whose issuer has subsequently gone into default. Bonds that trade without accrued interest are said to be trading *flat* or *clean*. By definition therefore,

clean price of a bond = dirty price − accrued interest

For bonds that are trading ex-dividend, the accrued coupon is negative and would be subtracted from the clean price. The calculation is given by (4.18):

$$AI = -C \quad \frac{(\text{days to } \textit{next} \text{ coupon})}{\textit{DayBase}} \tag{4.18}$$

Certain classes of bonds, for example US Treasuries and Eurobonds, do not have an *ex dividend* period and therefore trade *cum dividend* right up to the coupon date.

The accrued interest calculation for a bond is dependent on the day-count basis specified for the bond in question. When bonds are traded in the market the actual consideration that changes hands is made up of the clean price of the bond together with the accrued that has accumulated on the bond since the last coupon payment; these two components make up the dirty price of the bond. When calculating the accrued interest, the market will use the appropriate day-count convention for that bond. A particular market will apply one of five different methods to calculate accrued interest. These are:

actual/365	Accrued = Coupon × days/365
actual/360	Accrued = Coupon × days/360
actual/actual	Accrued = Coupon × days/actual number of days in the interest period
30/360	See below
30E/360	See below

When determining the number of days in between two dates, the first date is included but not the second; thus, under the actual/365 convention, there are 37 days between 4 August and 10 September. The last two conventions assume 30 days in each month, so for example there are '30 days' between 10 February and 10 March. Under the 30/360 convention, if the first date falls on the 31st, it is changed to the 30th of the month, and if the second date falls on the 31st *and*

the first date is the 30th or 31st, the second date is changed to the 30th. The difference under the 30E/360 method is that if the second date falls on the 31st of the month it is automatically changed to the 30th.

Example 4.3: Accrual calculation for US Treasury 3.625% August 2019

This US Treasury security has coupon dates of 15 February and 15 August each year. $100 nominal is traded for value on 15 September 2009. What is the accrued interest on the value date?

On value date 31 days have passed since last coupon date. The accrued interest calculation uses the actual number of days between the two coupon dates for the denominator, giving us:

$$3.625 \times 31/184 \times 0.5 = 0.305367$$

Example 4.4

Mansur buys £25,000 nominal of the 8% 2015 gilt for value on 27 August 1998, at a price of 102.4375. How much does he actually pay for the bond?

The clean price of the bond is 102.4375. The dirty price of the bond is:

$$102.4375 + 1.55342 = 103.99092$$

The total consideration is therefore $1.0399092 \times 25{,}000 =$ £25,997.73.

Example 4.5

A Norwegian government bond with a coupon of 8% is purchased for settlement on 30 July 1999 at a price of 99.50. Assume that this is seven days before the coupon date and therefore the bond trades ex-dividend. What is the all-in price?

The accrued interest is:

$$-8 \times \frac{7}{365} = -0.153424$$

The all-in price is therefore $99.50 - 0.1534 = 99.3466$.

Example 4.6

A bond has coupon payments on 1 June and 1 December each year. What is the day-base count if the bond is traded for value date on 30 October, 31 October and 1 November 1999 respectively? There are 183 days in the interest period.

	30 October	*31 October*	*1 November*
Act/365	151	152	153
Act/360	151	152	153
Act/Act	151	152	153
30/360	149	150	151
30E/360	149	150	150

Bond pricing and yield: the current approach

We are familiar with two types of fixed income security: *zero-coupon bonds*, also known as *discount bonds* or *strips*, and *coupon bonds*. A zero-coupon bond makes a single payment on its maturity date, while a coupon bond makes regular interest payments at regular dates up to and including its maturity date. A coupon bond may be regarded as a set of strips, with each coupon payment and the redemption payment on maturity being equivalent to a zero-coupon bond maturing on that date. The literature we review in this section is set in a market of default-free bonds, whether they are zero-coupon bonds or coupon bonds. The market is assumed to be liquid so that bonds may be freely bought and sold. Prices of bonds are determined by the economy-wide supply and demand for the bonds at any time, so they are *macroeconomic* and not set by individual bond issuers or traders.

Zero-coupon bonds

A zero-coupon bond is the simplest fixed income security. It is an issue of debt, the issuer promising to pay the face value of the debt to the bondholder on the date the bond matures. There are no coupon payments during the life of the bond, so it is a discount instrument, issued at a price that is below the face or principal amount. We denote as $P(t, T)$ the price of a discount bond at time t that matures at time T, with $T \geq t$. The term to maturity of the bond is denoted with n, where $n = T - t$. The price increases over time until the maturity date, when it reaches the maturity or *par* value. If the par value of the bond is £1, then the yield to maturity of the bond at time t is denoted by $r(t,T)$, where r is actually 'one plus the percentage yield' that is earned by holding the bond from t to T. We have:

$$P(t,T) = \frac{1}{[r(t,T)]^n} \tag{4.19}$$

The yield may be obtained from the bond price and is given by

$$r(t,T) = \left[\frac{1}{P(t,T)}\right]^{1/n} \tag{4.20}$$

which is sometimes written as

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

Analysts and researchers frequently work in terms of logarithms of yields and prices, or continuously compounded rates. One advantage of this is that it converts the non-linear relationship in (4.20) into a linear relationship.[12]

The bond price at time t_2 where $t \leq t_2 \leq T$ is given by:

$$P(t_2,T) = P(t,T)e^{(t_2-t)r(t,T)} \tag{4.22}$$

which is natural given that the bond price equation in continuous time is:

$$P(t,T) = e^{-r(t,T)(T-t)} \tag{4.23}$$

so that the yield is given by

$$r(t,T) = -\log\left(\frac{\mathrm{P(t,T)}}{\mathrm{n}}\right) \tag{4.24}$$

which is sometimes written as

$$\log r(t,T) = -\left(\frac{1}{n}\right)\log P(t,T) \tag{4.25}$$

The expression in (4.22) and (4.23) includes the exponential function, hence the use of the term continuously compounded.

The *term structure of interest rates* is the set of zero-coupon yields at time t for all bonds ranging in maturity from $(t, t+1)$ to $(t, t+m)$ where the bonds have maturities of $\{0,1,2,\ldots,m\}$. A good definition of the term structure of interest rates is given in Sundaresan, who states that it 'refers to the relationship between the

[12]A linear relationship in X would be a function $Y = f(X)$ in which the X values change via a power or index of 1 only and are not multiplied or divided by another variable or variables. So for example terms such as X^2, $\sqrt{X}$ and other similar functions are not linear in X, nor are terms such as XZ or X/Z where Z is another variable. In econometric analysis, if the value of Y is solely dependent on the value of X, then its rate of change with respect to X, or the derivative of Y with respect to X, denoted dY/dX, is independent of X. Therefore if $Y = 5X$, then $dY/dX = 5$, which is independent of the value of X. However if $Y = 5X^2$, then $dY/dX = 10X$, which is not independent of the value of X. Hence this function is not linear in X. The classic regression function $E(Y \mid X_i) = \alpha + \beta X_i$ is a linear function with slope β and intercept α and the regression 'curve' is represented geometrically by a straight line.

yield to maturity of default-free zero coupon securities and their maturities' (Sundaresan 1997, p. 176).

The *yield curve* is a plot of the set of yields for $r(t,t+1)$ to $r(t,t+m)$ against m at time t. For example, Figures 4.3 to 4.5 show the log zero-coupon yield curve for US Treasury strips, UK gilt strips and French OAT strips on 27 September 2000. Each of the curves exhibits peculiarities in its shape, although the most common type of curve is gently upward sloping, as is the French curve. The UK curve is *inverted*. We explore further the shape of the yield curve later in this chapter.

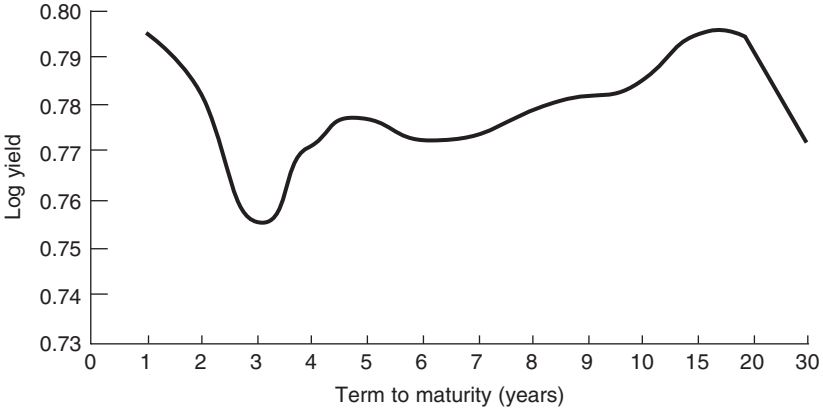


Figure 4.3 US Treasury zero-coupon yield curve, September 2000
Source: Bloomberg L.P.

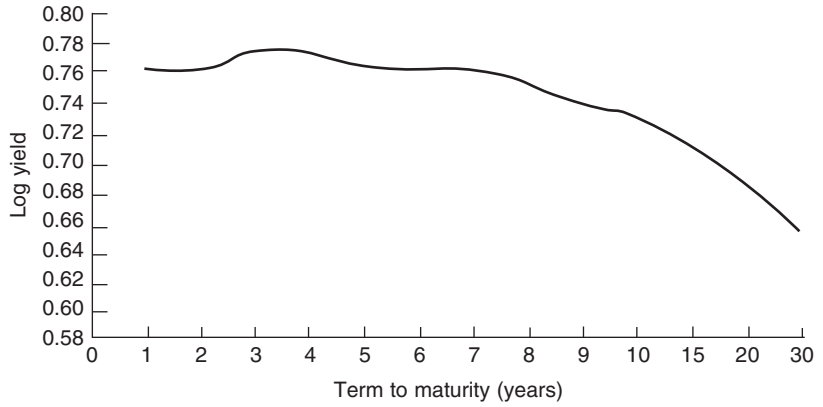


Figure 4.4 UK gilt zero-coupon yield curve, September 2000
Source: Bloomberg L.P.

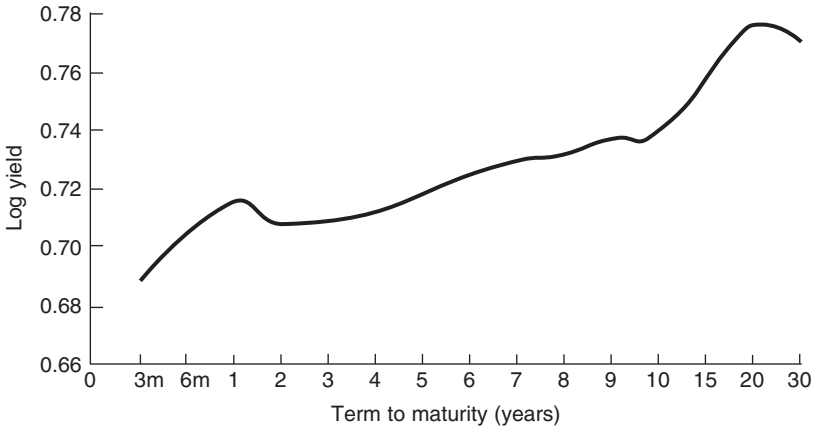


Figure 4.5 French OAT zero-coupon yield curve, September 2000
Source: Bloomberg L.P.

Coupon bonds

The majority of bonds in the market make periodic interest or *coupon* payments during their life, and are known as coupon bonds. We have already noted that such bonds may be viewed as a package of individual zero-coupon bonds. The coupons have a nominal value that is a percentage of the nominal value of the bond itself, with steadily longer maturity dates, while the final redemption payment has the nominal value of the bond itself and is redeemed on the maturity date. We denote a bond issued at time i and maturing at time T as having a w-element vector of payment dates $(t_1,t_2,\ldots t_{w-1},T)$ and matching date payments $C_1,C_2,\ldots C_{w-1},C_w$. In the academic literature these coupon payments are assumed to be made in continuous time, so that the stream of coupon payments is given by a positive function of time $C(t)$, $i < t \leq T$. An investor that purchases a bond at time t that matures at time T pays $P(t,T)$ and will receive the coupon payments as long as she continues to hold the bond.[13]

The yield to maturity at time t of a bond that matures at T is the interest rate that relates the price of the bond to the future returns on the bond, that is, the rate that discounts the bond's cash flow stream C_w to its price $P(t,T)$. This is given by

$$P(t,T) = \sum_{t_i > t} C_i e^{-(t_i - t)r(t,T)} \tag{4.26}$$

[13]In theoretical treatment this is the discounted clean price of the bond. For coupon bonds in practice, unless the bond is purchased for value on a coupon date, it will be traded with interest accrued. The interest that has accrued on a pro-rata basis from the last coupon date is added to the clean price of the bond, to give the market 'dirty' price that is actually paid by the purchaser.

which says that the bond price is given by the present value of the cash flow stream of the bond, discounted at the rate $r(t,T)$. For a zero-coupon bond (4.26) reduces to (4.24). In the academic literature where coupon payments are assumed to be made in continuous time, the Σ summation in (4.26) is replaced by the $\int$ integral. We will look at this in a moment.

In some texts the plot of the yield to maturity at time t for the term of the bonds m is described as the term structure of interest rates, but it is generally accepted that the term structure is the plot of zero-coupon rates only. Plotting yields to maturity is generally described as graphically depicting the yield curve, rather than the term structure. Of course, given the law of one price, there is a relationship between the yield to maturity yield curve and the zero-coupon term structure, and given the first, one can derive the second.

The expression at (4.26) obtains the continuously compounded yield to maturity $r(t,T)$. It is the use of the exponential function that enables us to describe the yield as continuously compounded.

The market frequently uses the measure known as *current yield* which is:

$$rc = \frac{C}{P_d} \times 100 \tag{4.27}$$

where P_d is the dirty price of the bond. The measure is also known as the *running yield* or *flat yield*. Current yield is not used to indicate the interest rate or discount rate and therefore should not be mistaken for the yield to maturity.

Bond price in continuous time[14]

Fundamental concepts

In this section we present an introduction to the bond price equation in continuous time. The necessary background on price processes is given in Choudhry (2001).

Consider a trading environment where bond prices evolve in a w-dimensional process

$$X(t) = [X_1(t), X_2(t), X_3(t),..., X_w(t)], \; t > 0 \tag{4.28}$$

where the random variables are termed *state variables* that reflect the state of the economy at any point in time. The markets assume that the state variables evolve

[14]This section follows the approach adopted in such texts as Avellaneda (2000), Baxter and Rennie (1996), Neftci (2000), Campbell *et al.* (1997), Ross (1999), and Shiller (1990). These are all excellent texts of very high quality, and strongly recommended. For an accessible and highly readable introduction Ross's book is worth buying for chapter 4 alone, as is Avellaneda's for his chapter 12. For a general introduction to the main pricing concepts see Campbell *et al.* (1997), chapter 10. Chapter 3 in Jarrow (1996) is an accessible introduction for discrete-time bond pricing. Sundaresan (1997) is an excellent overview text on the fixed income market as a whole, and is highly recommended. Further recommended references are given in the bibliography.

through a process described as geometric Brownian motion or a Weiner process. It is therefore possible to model the evolution of these variables, in the form of a stochastic differential equation.

The market assumes that the cash flow stream of assets such as bonds and (for equities) dividends is a function of the state variables. A bond is characterised by its coupon process:

$$C(t) = \mathrm{C}[X_1(t), X_2(t), X_3(t),..., X_w(t), t] \tag{4.29}$$

The coupon process represents the cash flow that the investor receives during the time that she holds the bond. Over a small incremental increase in time of dt from the time t the investor can purchase $1 + C(t)dt$ units of the bond at the end of the period $t + dt$. Assume that there is a very short-term discount security such as a Treasury bill that matures at $t + dt$, and during this period the investor receives a return of $r(t)$. This rate is the annualised short-term interest rate or *short rate*, which in the mathematical analysis is defined as the rate of interest charged on a loan that is taken out at time t and which matures almost immediately. For this reason the rate is also known as the *instantaneous rate*. The short rate is given by:[15]

$$r(t) = r(t,t) \tag{4.30}$$

and

$$r(t) = -\frac{\partial}{\partial T} \log P(t,t) \tag{4.31}$$

If we continuously reinvest the short-term security such as the T-bill at this short rate, we obtain a cumulative amount that is the original investment multiplied by (4.32).[16]

$$M(t) = \exp\left[\int_0^t r(s)ds\right] \tag{4.32}$$

[15]Generally the expression $r(t,T)$ is used to denote the zero-coupon interest rate starting at time t and maturing at time T, with $T > t$. Generally t is taken to be now, so that $t = 0$. The rate is the return on a risk-free zero-coupon bond of price P at time t and maturity T, so we may define it as $P(t,T) = 1/[1 + r(t,T)^T]$. The term $r(t)$ is generally used to express the spot interest rate at the limit of $r(t,T)$ as T approaches t. Thus $r(t)$ may be regarded as the continuously compounded rate of return on a risk-free zero-coupon bond of infinitesimal maturity. The spot rate is therefore a theoretical construct, as it is unlikely to be observed directly on a market instrument. The zero-coupon rate on an instrument of infinitesimal maturity is sometimes expressed as $r(t, t)$ or as we do here, $r(t)$. So at time $t = 0$ we have $r(t) = \lim_{T \to 0} r(t,T)$.

[16]This expression uses the integral operator. The integral is the tool used in mathematics to calculate sums of an infinite number of objects, that is where the objects are uncountable. This is different from the Σ operator which is used for a countable number of objects. For a readable and accessible review of the integral and its use in quantitative finance, see Neftci (2000), pp. 59–66, a review of which is given in appendix 3.1 of Choudhry (2001).

where M is a money market account that offers a return of the short rate $r(t)$.

If we say that the short rate is constant, making $r(t) = r$, then the price of a risk-free bond that pays £1 on maturity at time T is given by

$$P(t,T) = e^{-r(T-t)} \tag{4.33}$$

What (4.33) states is that the bond price is simply a function of the continuously-compounded interest rate, with the right-hand side of (4.33) being the discount factor at time t. At $t = T$ the discount factor will be 1, which is the redemption value of the bond and hence the price of the bond at this time.

Consider the following scenario. A market participant may undertake the following:

- It can invest $e^{-r(T-t)}$ units cash in a money market account today, which will have grown to a sum of £1 at time T.
- It can purchase the risk-free zero-coupon bond today, which has a maturity value of £1 at time T.

The market participant can invest in either instrument, both of which we know beforehand to be risk-free, and both of which have identical payouts at time T and have no cash flow between now and time T. As interest rates are constant, a bond that paid out £1 at T must have the same value as the initial investment in the money market account, which is $e_1{}^{-r(T-t)}$. Therefore equation (4.33) must apply. This is a restriction placed on the zero-coupon bond price by the requirement for markets to be arbitrage-free.

If the bond was not priced at this level, arbitrage opportunities would present themselves. Consider if the bond was priced higher than $e_t{}^{-r(T-t)}$. In this case, an investor could sell short the bond and invest the sale proceeds in the money market account. On maturity at time T, the short position will have a value of −£1 (negative, because the investor is short the bond) while the money market will have accumulated £1, which the investor can use to pay the proceeds on the zero-coupon bond. However the investor will have surplus funds because at time *t:*

$$P(t,T) = e^{-r(T-t)} > 0$$

and so will have profited from the transaction at no risk to himself.

The same applies if the bond is priced below $e_1{}^{-r(T-t)}$. In the case the investor borrows $e_i{}^{-r(T-t)}$ and buys the bond at its price $P(t,T)$. On maturity the bond pays £1 which is used to repay the loan amount. However the investor will gain because:

$$e^{-r(T-t)} - P(t,T) > 0$$

Therefore the only price at which no arbitrage profit can be made is if

$$P(t,T) = e^{-r(T-t)} \tag{4.34}$$

In the academic literature the price of a zero-coupon bond is given in terms of the evolution of the short-term interest rate, in what is termed the *risk-neutral measure*.[17] The short rate $r(t)$ is the interest rate earned on a money market account or short-dated risk-free security such as the T-bill suggested above, and it is assumed to be continuously compounded. This makes the mathematical treatment simpler. With a zero-coupon bond we assume a payment on maturity of 1 (say \$1 or £1), a one-off cash flow payable on maturity at time T. The value of the zero-coupon bond at time t is therefore given by

$$P(t,T) = \exp\left[-\int_t^{Tr(s)ds}\right] \tag{4.35}$$

which is the redemption value of 1 divided by the value of the money market account, given by (4.32).

The bond price for a coupon bond is given in terms of its yield as:

$$P(t, T) = \exp(-(T - t)r(T - t)) \tag{4.36}$$

Expression (4.35) is very commonly encountered in the academic literature. Its derivation is not so frequently occurring however; readers will find it in Ross (1999). This reference is highly recommended reading. It is also worth referring to Neftci (2000), chapter 18.

The expression (4.35) represents the zero-coupon bond pricing formula when the spot rate is continuous or stochastic, rather than constant. The rate $r(s)$ is the risk-free return earned during the very short or *infinitesimal* time interval $(t, t + dt)$. The rate is used in the expressions for the value of a money market account (4.32) and the price of a risk-free zero-coupon bond (4.36).

Stochastic rates in continuous time

In the academic literature the bond price given by (4.36) evolves as a *martingale* process under the risk-neutral probability measure $\tilde{P}$. This is an advanced branch of fixed income mathematics, and is outside the scope of this book. (However it will be introduced in introductory fashion in the next chapter.[18]) However under this analysis the bond price is given as:

$$P(t, T) = E_t^{\tilde{P}}\left[\mathrm{e}^{-\int_t^T r(s)ds}\right] \tag{4.37}$$

where the right-hand side of (4.37) is viewed as the randomly evolved *discount factor* used to obtain the present value of the £1 maturity amount. Expression (4.37) also

[17]This is part of the *arbitrage pricing theory*. For detail on this see Cox *et al.* (1985), while Duffie (1992) is a fuller treatment for those with a strong grounding in mathematics.

[18]Interested readers should consult Nefcti (2000), chapters 2 and 17–18. Another accessible text is Baxter and Rennie (1996), while Duffie (1992) is a leading reference for those with a strong background in financial mathematics.

states that bond prices are dependent on the entire spectrum of short-term interest rates $r(s)$ in the future during the period $t < s < T$. This also implies that the term structure at time t contains all the information available on short rates in the future.[19]

From (4.37) we say that the function $T \to P_t^T, t < T$ is the discount curve (or *discount function*) at time t. Avellaneda (2000) notes that the markets usually replace the term $(T - t)$ with a term meaning *time to maturity,* so the function becomes

$$\tau \to P_t^{T+\tau}, \tau > 0 \text{ where, } \tau = (T - t).$$

Under a constant spot rate, the zero-coupon bond price is given by:

$$P(t,T) = e^{-r(t,T)(T-t)} \tag{4.38}$$

From (4.37) and (4.38) we can derive a relationship between the yield $r(t,T)$ of the zero-coupon bond and the short rate $r(t)$, if we equate the two right-hand sides, namely:

$$e^{-r(t,T)(T-t)} = E_t^{\tilde{P}}\left[\mathrm{e} - \int_t^T r(s)ds\right] \tag{4.39}$$

Taking the logarithm of both sides we obtain

$$r(t,T) = \frac{-\log E_t^{\tilde{P}}\left[\mathrm{e} - \int_t^T r(s)ds\right]}{T-t} \tag{4.40}$$

This describes the yield on a bond as the average of the spot rates that apply during the life of the bond, and under a constant spot rate the yield is equal to the spot rate.

With a zero-coupon bond and assuming that interest rates are positive, $P(t,T)$ is less than or equal to 1. The yield of the bond is, as we have noted, the continuously compounded interest rate that equates the bond price to the discounted present value of the bond at time t. This is given by

$$r(t,T) = -\frac{\log(P(t,T))}{T-t} \tag{4.41}$$

so we obtain

$$P(t,T) = e^{-(T-t)r(T-t)} \tag{4.42}$$

In practice this means that an investor will earn $r(t,T)$ if she purchases the bond at t and holds it to maturity.

[19]This is related to the view of the short rate evolving as a martingale process. For a derivation of (4.37) see Neftci (2000), p. 417.

Coupon bonds

Using the same principles as in the previous section, we can derive an expression for the price of a coupon bond in the same terms of a risk-neutral probability measure of the evolution of interest rates. Under this analysis, the bond price is given by:

$$P_c = 100.E_t^{\tilde{P}}\left(e^{-\int_s^{t_N} r(s)ds}\right) + \sum_{nt_n>t}^{N} \frac{C}{w} E_t^{\tilde{P}}\left(e^{-\int_t^{t_n} r(s)ds}\right) \tag{4.43}$$

where

P_c is the price of a coupon bond
C is the bond coupon
t_n is the coupon date, with $n \leq N$, and $t = 0$ at the time of valuation
w is the coupon frequency.[20]

and where 100 is used as the convention for principal or bond nominal value (that is, prices are quoted%, or per 100 nominal).

Expression (4.43) is written in some texts as:

$$P_c = 100e^{-rN} + \int_n^N Ce^{-r^2}dt \tag{4.44}$$

We can simplify (4.43) by substituting Df to denote the discount factor part of the expression and assuming an annual coupon, which gives us:

$$P = 100Df_N + \sum_{nt_n>t}^{N} CDf_N \tag{4.45}$$

This states that the market value of a risk-free bond on any date is determined by the discount function on that date.

We know that the actual price paid in the market for a bond includes accrued interest from the last coupon date, so that the price given by (4.45) is known as the clean price and the traded price, which includes accrued interest, is known as the *dirty price*.

Forward rates

An investor can combine positions in bonds of differing maturities to guarantee a rate of return that begins at a point in the future. That is, the trade ticket would be written at time t but would cover the period T to $T + 1$ where $t < T$ (sometimes written as beginning at T_1 and ending at T_2, with $t < T_1 < T_2$). The interest rate earned during this period is known as the *forward rate*.[21] The mechanism by which

[20] Conventional or *plain vanilla* bonds pay coupon on an annual or semi-annual basis. Other bonds, notably certain floating-rate notes and mortgage and other asset-backed securities, also pay coupon on a monthly basis, depending on the structuring of the transaction.

[21] See the footnote on p. 639 of Shiller (1990) for a fascinating insight on the origin of the term 'forward rate', which Shiller ascribes to John Hicks in his book *Value and Capital* (1946).

The forward rate

An investor buys at time t one unit of a zero-coupon bond maturing at time T, priced at $P(t, T)$ and simultaneously sells $P(t,T)/P(t,T+1)$ bonds that mature at $T+1$. From the table below we see that the net result of these transactions is a zero cash flow. At time T there is a cash inflow of 1, and then at time $T + 1$ there is a cash outflow of $P(t,T)/P(t,T+1)$. These cash flows are identical to a loan of funds made during the period T to $T+1$, contracted at time t. The interest rate on this loan is given by $P(t,T)/P(t,T+1)$, which is therefore the forward rate. That is,

$$f(t,T) = \frac{P(t,T)}{P(t,T+1)} \tag{4.46}$$

Together with our earlier relationships on bond price and yield, from (4.46) we can define the forward rate in terms of yield, with the return earned during the period $(T,T+1)$ being:

$$f(t,T+1) = \frac{1}{(P(t,T+1)/P(t,T))} = \frac{(r(t,T+1))^{(T+1)}}{r(t,T)^T} \tag{4.47}$$

Transactions	*Time*		
	t	*T*	*T+1*
Buy 1 unit of T-period bond	−P(t,T)	+1	
Sell P(t,T)/P(t, T+1) T+1 period bonds	+[(P(t,T)/P(t, T+1)]P(t, T+1)		−P(t,T)/P(t, T+1)
Net cash flows	0	+1	−P(t,T)/P(t, T+1)

From (4.46) we can obtain a bond price equation in terms of the forward rates that hold from t to T,

$$P(t,T) = \frac{1}{\Pi_{k=t}^{T-1} f(t,k)} \tag{4.48}$$

A derivation of this expression can be found in Jarrow (1996), chapter 3. Equation (4.48) states that the price of a zero-coupon bond is equal to the nominal value, here assumed to be 1, receivable at time T after it has been discounted at the set of forward rates that apply from t to T.[22]

[22]The symbol $\prod$ means 'take the product of', and is defined as $\prod_{t=1}^{n} x_i = x_1 \cdot x_2 \cdot \ldots \cdot x_n$, so that $\prod_{k=t}^{T-1} f(t,k) = f(t,t) \cdot f(t,t+1) \cdot \ldots \cdot f(t,T-1)n$ which is the result of multiplying the rates that are obtained when the index k runs from t to $T - 1$.

this forward rate can be guaranteed is described in the box opposite, following Jarrow (1996) and Campbell, Lo and McKinlay (1997).

When calculating a forward rate, remember that it is as if we are writing an interest rate contract today that comes into effect at a start date some time in the *future*. In other words we are trading a forward contract. The law of one price, or no-arbitrage, is used to calculate the rate. For a loan that begins at T and matures at $T+1$, similarly to the way we described in the box, consider a purchase of a $T+1$ period bond and a sale of p amount of the T-period bond. The cash net cash position at t must be zero, so p is given by

$$p = \frac{P(t,T+1)}{P(t,T)} \tag{4.49}$$

and to avoid arbitrage the value of p must be the price of the $T+1$-period bond at time T. Therefore the forward yield is given by:

$$f(t,T+1) = \frac{\log P(t,T+1) - \log P(t,T)}{(T+1)-T} \tag{4.50}$$

If the period between T and the maturity of the later-dated bond is reduced, so we now have bonds that mature at T and T_2, and $T_2 = T + \Delta t$, then as the incremental change in time Δt becomes progressively smaller we eventually obtain an instantaneous forward rate, which is given by

$$f(t,T) = -\frac{\partial}{\partial T}\log P(t,T) \tag{4.51}$$

This rate is defined as the forward rate and is the price today of forward borrowing at time T. The forward rate for borrowing today where $T = t$ is equal to the instantaneous short rate $r(t)$. At time t the spot and forward rates for the period (t,t) will be identical, at other maturity terms they will differ.

For all points other that at (t,t) the forward rate yield curve will lie above the spot rate curve if the spot curve is positively sloping. The opposite applies if the spot rate curve is downward sloping. Campbell *et al.* (1997, pp. 400–1) observe that this property is a standard one for marginal and average cost curves. That is, when the cost of a marginal unit (say, of production) is above that of an average unit, then the average cost will increase with the addition of a marginal unit. This results in the average cost rising when the marginal cost is above the average cost. Equally the average cost per unit will decrease when the marginal cost lies below the average cost.

The term structure of interest rates

We have already referred to the yield curve or *term structure of interest rates*. Strictly speaking only a spot rate yield curve is a term structure, but one sometimes encounters the two expressions being used synonymously. At any time t there will

be a set of coupon and/or zero-coupon bonds with different terms to maturity and cash flow streams. There will be certain fixed maturities that are not represented by actual bonds in the market, as there will be more than one bond maturing at or around the same redemption date. The debt capital markets and the pricing of debt instruments revolve around the term structure, and for this reason this area has been extensively researched in the academic literature. There are a number of ways to estimate and interpret the term structure, and in this section we review the bootstrapping technique using coupon bond yields, which follows the approach described in Windas (1993).

Spot and forward rates: bootstrapping from coupon bond yields and the par yield curve

Par, spot and forward rates have a close mathematical relationship. Here we explain and derive these different interest rates and explain their application in the markets. Note that *spot* interest rates are also called *zero-coupon* rates, because they are the interest rates that would be applicable to a zero-coupon bond. The two terms are used synonymously, however strictly speaking they are not identical. Zero-coupon bonds are actual market instruments, and the yield on zero-coupon bonds can be observed in the market. A spot rate is a purely theoretical construct, and so cannot actually be observed directly. For our purposes though, we will use the terms synonymously.

A *par* yield is the yield to maturity on a bond that is trading at par. This means that the yield is equal to the bond's coupon level. A zero-coupon bond is a bond that has no coupons, and therefore only one cash flow; the redemption payment on maturity. It is thus a *discount* instrument, as it is issued at a discount to par and redeemed at par. The yield on a zero-coupon bond can be viewed as a true yield, at the time that it is purchased, if the paper is held to maturity. This is because no reinvestment of coupons is involved and so there are no interim cash flows vulnerable to a change in interest rates. Zero-coupon yields are the key determinant of value in the capital markets, and they are calculated and quoted for every major currency. Zero-coupon rates can be used to value any cash flow that occurs at a future date. Where zero-coupon bonds are traded, the yield on a zero-coupon bond of a particular maturity is the zero-coupon rate for that maturity. Not all debt capital trading environments possess a liquid market in zero-coupon bonds.

However it is not necessary to have zero-coupon bonds in order to calculate zero-coupon rates. It is possible to calculate zero-coupon rates from a range of market rates and prices, including coupon bond yields, interest-rate futures and currency deposits.

We describe here the close mathematical relationship between par, zero-coupon and forward rates. We also illustrate how the bootstrapping technique can be used to calculate spot and forward rates from coupon bond redemption yields. In addition, once the discount factors are known, any of these rates can be calculated. The relationship between the three rates allows the markets to price interest-rate swap and FRA rates, as a swap rate is the weighted arithmetic average of forward rates for the term in question.

Implied spot and forward rates

We describe here how to obtain zero-coupon and forward interest rates from the yields available from coupon bonds, using a method known as *bootstrapping*. In a government bond market such as that for US Treasuries or UK gilts, the bonds are considered to be *default-free*. The rates from a government bond yield curve describe the risk-free rates of return available in the market *today*, however they also *imply* (risk-free) rates of return for *future time periods*. These implied future rates, known as *implied forward rates*, or simply *forward rates*, can be derived from a given spot yield curve using bootstrapping. This term reflects the fact that each calculated spot rate is used to determine the next period spot rate, in successive steps.

Table 4.4 shows a hypothetical benchmark gilt yield curve for value as at 7 December 2000, with the bonds all priced at par. The observed yields of the benchmark bonds that compose the curve are displayed in the last column. All rates are annualised and assume semi-annual compounding. The bonds all pay on the same coupon dates of 7 June and 7 December, and as the value date is a coupon date, there is no accrued interest on any of the bonds.[23] The *clean* and *dirty* prices for each bond are identical.

Table 4.4 Hypothetical UK government bond yields as on 7 December 2000

Bond	*Term to maturity (years)*	*Coupon*	*Maturity date*	*Price*	*Gross Redemption Yield*
4% Treasury 2001	0.5	4%	7-Jun-01	100	4%
5% Treasury 2001	1	5%	7-Dec-01	100	5%
6% Treasury 2002	1.5	6%	7-Jun-02	100	6%
7% Treasury 2002	2	7%	7-Dec-02	100	7%
8% Treasury 2003	2.5	8%	7-Jun-03	100	8%
9% Treasury 2003	3	9%	7-Dec-03	100	9%

The gross redemption yield or *yield to maturity* of a coupon bond describes the single rate that present-values the sum of all its future cash flows to its current price. It is essentially the *internal rate of return* of the set of cash flows that make up the bond. This yield measure suffers from a fundamental weakness in that each cash flow is present-valued at the same rate, an unrealistic assumption in anything other than a flat yield curve environment. So the yield to maturity is an *anticipated* measure of the return that can be expected from holding the bond from purchase until maturity. In practice it will only be achieved under the following conditions:

- the bond is purchased on issue
- all the coupons paid throughout the bond's life are reinvested at the same yield to maturity at which the bond was purchased
- the bond is held until maturity

[23] Benchmark gilts pay coupon on a semi-annual basis on 7 June and 7 December each year.

In practice these conditions will not be fulfilled, and so the yield to maturity of a bond is not a true interest rate for that bond's maturity period.

The bonds in Table 4.4 pay semi-annual coupons on 7 June and 7 December and have the same time period – six months – between 7 December 2000 (their valuation date) and 7 June 2001 (their next coupon date). However since each issue carries a different yield, the next six-month coupon payment for each bond is present-valued at a different rate. In other words, the six-month bond present-values its six-month coupon payment at its 4% yield to maturity, the one-year at 5%, and so on. Because each of these issues uses a different rate to present-value a cash flow occurring at the same future point in time, it is unclear which of the rates should be regarded as the true interest rate or benchmark rate for the six-month period from 7 December 2000 to 7 June 2001. This problem is repeated for all other maturities.

For the purposes of valuation and analysis however, we require a set of true interest rates, and so these must be derived from the redemption yields that we can observe from the benchmark bonds trading in the market. These rates we designate as rs_i, where rs_i is the *implied spot rate* or *zero-coupon rate* for the term beginning on 7 December 2000 and ending at the end of period i.

We begin calculating implied spot rates by noting that the six-month bond contains only one future cash flow, the final coupon payment and the redemption payment on maturity. This means that the bond is in effect trading as a zero-coupon bond, as there is only one cash flow left (the final payment). Since this cash flow's present value, future value and maturity term are known, the unique interest rate that relates these quantities can be solved using the compound interest equation (4.52) below.

$$\text{FV} = \text{PV} \times \left(1 + \frac{rs_i}{m}\right)^{(nm)} \tag{4.52}$$

$$rs_i = m \times \left(\sqrt[(nm)]{\frac{\text{FV}}{\text{PV}}} - 1 \right)$$

where

FV is the future value
PV is the present value
rs_i is the implied i-period spot rate
m is the number of interest periods per year
n is the number of years in the term

The first rate to be solved is referred to as the implied six-month spot rate and is the true interest rate for the six-month term beginning on 2 January and ending on 2 July 2000.

Equation (4.52) relates a cash flow's present value and future value in terms of an associated interest rate, compounding convention and time period. Of course if we re-arrange it, we may use it to solve for an implied spot rate. For the six-month bond the final cash flow on maturity is £102, comprised of the £2 coupon payment and the £100 (*par*) redemption amount. So we have for the first term, $i = 1$, FV = £102, PV = £100, $n = 0.5$ years and $m = 2$. This allows us to calculate the spot rate as follows:

$$\begin{aligned} rs_i &= m \times \left(\sqrt[(nm)]{\text{FV/PV}} - 1 \right) \\ rs_1 &= 2 \times \left(\sqrt[(0.5 \times 2)]{\text{£102/£100}} - 1 \right) \\ rs_1 &= 0.04000 \\ rs_1 &= 4.000\% \end{aligned} \tag{4.53}$$

Thus the implied six-month spot rate or zero-coupon rate is equal to 4%.[24] We now need to determine the implied one-year spot rate for the term from 7 December 2000 to 7 June 2001. We note that the one-year issue has a 5% coupon and contains two future cash flows: a £2.50 six-month coupon payment on 7 June 2001 and a £102.50 one-year coupon and principal payment on 7 December 2001. Since the first cash flow occurs on 7 June – six months from now – it must be present-valued at the 4% six-month spot rate established above. Once this present value is determined, it may be subtracted from the £100 total present value (its current price) of the one-year issue to obtain the present value of the one-year coupon and cash flow. Again we then have a single cash flow with a known present value, future value and term. The rate that equates these quantities is the implied one-year spot rate. From equation (4) the present value of the six-month £2.50 coupon payment of the one-year benchmark bond, discounted at the implied six-month spot rate, is:

$$\begin{aligned} \text{PV}_{\text{6-mo cash flow, 1-yr bond}} &= \text{£2.50}/(1 + 0.04/2)^{(0.5 \times 2)} \\ &= \text{£2.45098} \end{aligned}$$

The present value of the one-year £102.50 coupon and principal payment is found by subtracting the present value of the six-month cash flow, determined above, from the total present value (current price) of the issue:

$$\begin{aligned} \text{PV}_{\text{1-yr cash flow, 1-yr bond}} &= \text{£100} - \text{£2.45098} \\ &= \text{£97.54902} \end{aligned}$$

[24] Of course intuitively we could have concluded that the six-month spot rate was 4%, without the need to apply the arithmetic, as we had already assumed that the six-month bond was a quasi-zero-coupon bond.

The implied one-year spot rate is then determined by using the £97.54902 present value of the one-year cash flow determined above:

$$rs_2 = 2 \times \left(\sqrt[(1\times2)]{£102.50/£97.54902} - 1 \right)$$
$$= 0.0501256$$
$$= 5.01256\%$$

The implied 1.5 year spot rate is solved in the same way:

$$PV_{\text{6-mo cash flow, 1.5-yr bond}} = £3.00/(1 + 0.04/2)^{(0.5\times2)}$$
$$= £2.94118$$

$$PV_{\text{1-yr cash flow, 1.5-yr bond}} = £3.00/(1 + 0.0501256/2)^{(1\times2)}$$
$$= £2.85509$$

$$PV_{\text{1.5-yr cash flow, 1.5-yr bond}} = £100 - £2.94118 - £2.85509$$
$$= £94.20373$$

$$rs_3 = 2 \times \left(\sqrt[(1.5\times2)]{£103 / £94.20373} - 1 \right)$$
$$= 0.0604071$$
$$= 6.04071\%$$

Extending the same process for the two-year bond, we calculate the implied two-year spot rate rs_4 to be 7.0906%. The implied 2.5-year and 3-year spot rates rs_5 and rs_6 are 8.1709% and 9.2879% respectively.

The interest rates rs_1, rs_2, rs_3, rs_4, rs_5 and rs_6 describe the true zero-coupon interest rates for the 6-month, 1-year, 1.5-year, 2-year, 2.5-year and 3-year terms that begin on 7 December 2000 and end on 7 June 2001, 7 December 2001, 7 June 2002, 7 December 2002, 7 June 2003 and 7 December 2003 respectively. They are also called *implied spot rates* because they have been calculated from redemption yields observed in the market from the benchmark government bonds that were listed in Table 4.4.

Note that the 1-, 1.5-, 2-, 2.5- and 3-year implied spot rates are progressively greater than the corresponding redemption yields for these terms. This is an important result, and occurs whenever the yield curve is positively sloped. The reason for this is that the present values of a bond's shorter-dated cash flows are discounted at rates that are lower than the bond's redemption yield; this generates higher present values that, when subtracted from the current price of the bond, produce a lower present value for the final cash flow. This lower present value implies a spot rate that is greater than the issue's yield. In an inverted yield curve environment we observe the opposite result; that is, implied rates that lie below the corresponding redemption yields. If the

redemption yield curve is flat, the implied spot rates will be equal to the corresponding redemption yields.

Once we have calculated the spot or zero-coupon rates for the 6-month, 1-year, 1.5-year, 2-year, 2.5-year and 3-year terms, we can determine the rate of return that is implied by the yield curve for the sequence of six-month periods beginning on 7 December 2000, 7 June 2001, 7 December 2001, 7 June 2002 and 7 December 2002. These period rates are referred to as *implied forward rates* or *forward-forward rates* and we denote these as rf_i, where rf_i is the implied six-month forward interest rate today for the ith period.

Since the implied six-month zero-coupon rate (spot rate) describes the return for a term that coincides precisely with the first of the series of six-month periods, this rate describes the risk-free rate of return for the first six-month period. It is therefore equal to the first period spot rate. Thus we have $rf_1 = rs_1 = 4.0\%$, where rf_1 is the risk-free forward rate for the first six-month period beginning at period 1. We show now how the risk-free rates for the second, third, fourth, fifth and sixth six-month periods, designated rf_2, rf_3, rf_4, rf_5 and rf_6 respectively may be solved from the implied spot rates.

The benchmark rate for the second semi-annual period rf_2 is referred to as the one-period forward six-month rate, because it goes into effect one six-month period from now ('one-period forward') and remains in effect for six months ('six-month rate'). It is therefore the six-month rate in six months' time, and is also referred to as the six-month forward-forward rate. This rate, in conjunction with the rate from the first period rf_1, must provide returns that match those generated by the implied one-year spot rate for the entire one-year term. In other words, one pound invested for six months from 7 December 2000 to 7 June 2001 at the first period's benchmark rate of 4% and then reinvested for another six months from 7 June 2001 to 7 December 2001 at the second period's (as yet unknown) implied *forward* rate must enjoy the same returns as one pound invested for one year from 7 December 2000 to 7 December 2001 at the implied one-year *spot* rate of 5.0125%. This reflects the law of no-arbitrage.

A moment's thought will convince us that this must be so. If this were not the case, we might observe an interest rate environment in which the return received by an investor over any given term would depend on whether an investment is made at the start period for the entire maturity term or over a succession of periods within the whole term and reinvested. If there were any discrepancies between the returns received from each approach, there would exist an unrealistic arbitrage opportunity, in which investments for a given term carrying a lower return might be sold short against the simultaneous purchase of investments for the same period carrying a higher return, thereby locking in a risk-free, cost-free profit. Therefore, forward interest rates must be calculated so that they are *arbitrage-free*. Forward rates are thus not a prediction of what spot interest rates are likely to be in the future, but rather a mathematically derived set of interest rates that reflect the current spot term structure and the rules of no-arbitrage. Excellent mathematical explanations of the no-arbitrage property of interest-rate markets can be found in Ingersoll (1987), Jarrow (1996), and Robert Shiller (1990) among others.

The existence of a no-arbitrage market of course simplifies the calculation of forward rates; we know that the return from an investment made over a period must equal the return made from investing in a shorter period and successively reinvesting to a matching term. If we know the return over the shorter period, we are left with only one unknown, the full-period forward rate, which is then easily calculated. In our example, having established the rate for the first six-month period, the rate for the second six-month period – the one-period forward six-month rate – is determined below.

The future value of £1 invested at rf_1, the period 1 forward rate, at the end of the first six-month period is calculated as follows:

$$\begin{aligned} \text{FV}_1 &= £1 \times \left(1 + \frac{rf_1}{2}\right)^{(0.5\times 2)} \\ &= £1 \times \left(1 + \frac{0.04}{2}\right)^{1} \\ &= £1.02000 \end{aligned}$$

The future value of £1 at the end of the one-year term, invested at the implied benchmark one-year spot rate, is determined as follows:

$$\begin{aligned} \text{FV}_2 &= £1 \times \left(1 + \frac{rs_2}{2}\right)^{(1\times 2)} \\ &= £1 \times \left(1 + \frac{0.0501256}{2}\right)^{2} \\ &= £1.050754 \end{aligned}$$

The implied benchmark one-period forward rate rf_2 is the rate that equates the value of FV_1 (£1.02) on 7 June 2001 to FV_2 (£1.050754) on 7 December 2001. From equation (4) we have:

$$\begin{aligned} rf_2 &= 2 \times \left(\sqrt[(0.5\times 2)]{\frac{\text{FV}_2}{\text{FV}_1}} - 1\right) \\ &= 2 \times \left(\frac{£1.050754}{£1.02} - 1\right) \\ &= 0.060302 \\ &= 6.0302\% \end{aligned}$$

In other words £1 invested from 7 December to 7 June at 4.0% (the implied forward rate for the first period) and then reinvested from 7 June to 7 December 2001 at 6.0302% (the implied forward rate for the second period) would accumulate the same returns as £1 invested from 7 December 2000 to 7 December 2001 at 5.01256% (the implied one-year spot rate).

The rate for the third six-month period – the two-period forward six-month interest rate – may be calculated in the same way:

$$\begin{aligned}
FV_2 &= £1.050754 \\
FV_3 &= £1 \times (1 + rs_3/2)^{(1.5\times 2)} \\
&= £1 \times (1 + 0.0604071/2)^3 \\
&= £1.093375
\end{aligned}$$

$$\begin{aligned}
rf_3 &= 2 \times \left(\sqrt[(0.5\times 2)]{\frac{FV_3}{FV_4}} - 1 \right) \\
&= 2 \times \left(\sqrt[1]{£1.093375 / £1.050754} - 1 \right) \\
&= 0.081125 \\
&= 8.1125\%
\end{aligned}$$

In the same way the three-period forward six-month rate rf_4 is calculated to be 10.27247%. The rest of the results are shown in Table 4.5. We say *one-period* forward rate because it is the forward rate that applies to the six-month period. The results of the implied spot (zero-coupon) and forward rate calculations along with the given redemption yield curve are illustrated graphically in Figure 4.6.

The simple bootstrapping methodology can be applied using a spreadsheet for actual market redemption yields. However in practice we will not have a set of bonds with exact and/or equal periods to maturity and coupons falling on the same date. Nor will they all be priced conveniently at par. In designing a spreadsheet spot rate calculator therefore, the coupon rate and maturity date is entered as standing data and interpolation is used when calculating the spot rates for bonds with uneven maturity dates. A spot curve model that uses this approach in conjunction with the bootstrapping method is available for downloading at www.yieldcurve.com. Market practitioners usually use discount factors to extract spot and forward rates from market prices. For an account of this method, see chapter 9. Chapter 11 describes using the Internet resource QuantLib via an Excel front end.

Table 4.5 Implied spot and forward rates

Term to maturity	*Yield to maturity*	*Implied spot rate*	*Implied one-period forward rate*
0.5	4.0000%	4.00000%	4.00000%
1	5.0000%	5.01256%	6.03023%
1.5	6.0000%	6.04071%	8.11251%
2	7.0000%	7.09062%	10.27247%
2.5	8.0000%	8.17090%	12.24833%
3	9.0000%	9.28792%	14.55654%

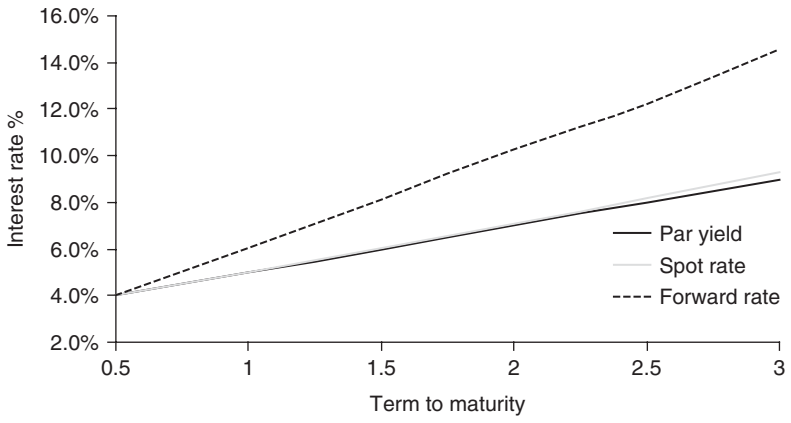


Figure 4.6 Par, spot and forward yield curves

Example 4.7

Consider the following spot yields:

one-year 10%
two-year 12%

Assume that a bank's client wishes to lock in *today* the cost of borrowing 1-year funds in one year's time. The solution for the bank (and the mechanism to enable the bank to quote a price to the client) involves raising 1-year funds at 10% and investing the proceeds for two years at 12%. As we observed above, the no-arbitrage principle means that the same return must be generated from both fixed rate and reinvestment strategies.

Using the following formula:

$$(1+rs_2)^2 = (1+rs_1)(1+rf)$$

$$rf = \frac{(1+y_2)^2}{(1+y_1)} - 1$$

the relevant forward rate is calculated to be 14.04%.

Example 4.8

If a one-year AAA Eurobond trading at par yields 10% and a two-year Eurobond of similar credit quality, also trading at par, yields 8.75%, what should be the price of a two-year AAA zero-coupon bond? Note that Eurobonds pay coupon annually.

(a)	Cost of two-year bond (% nominal)	100
(b)	*less* amount receivable from sale of first coupon on this bond (that is, its present value)	= 8.75 / 1 + 0.10 = 7.95
(c)	*equals* amount that must be received on sale of second coupon plus principal in order to break even	92.05
(d)	calculate the yield implied in the cash flows below (that is, the two-year zero-coupon yield) – receive 92.05 – pay out on maturity 108.75	
	Therefore Gives R equal to 8.69%	$92.05 = 108.75/(1 + R)^2$
(e)	What is the price of a two-year zero-coupon bond with nominal value 100, to yield 8.69%?	= (92.05/108.75) × 100 = 84.64

Example 4.9

A highly rated customer asks you to fix a yield at which he could issue a 2-year zero-coupon USD Eurobond in three years' time. At this time the US Treasury zero-coupon rates were:

1 Yr	6.25%
2 Yr	6.75%
3 Yr	7.00%
4 Yr	7.125%
5 Yr	7.25%

(a) Ignoring borrowing spreads over these benchmark yields, as a market maker you could cover the exposure created by borrowing funds for five years on a zero-coupon basis and placing these funds in the market for three years before lending them on to your client. Assume annual interest compounding (even if none is actually paid out during the life of the loans)

$$\text{Borrowing rate for 5 years } \left[\frac{R_5}{100}\right] = 0.0725$$

$$\text{Lending rate for 3 years } \left[\frac{R_3}{100}\right] = 0.0700$$

(b) The key arbitrage relationship is:

Total cost of funding = Total Return on Investments

$$(1 + R_5)^5 = (1 + R_3)^3 \times (1 + R_{3\times5})^2$$

Therefore the break-even forward yield is:

$$R_{3\times5} = \sqrt[2]{\left[\frac{(1+R_5)^5}{(1+R_3)^3}\right]} - 1$$

$$= 7.63\%$$

Example 4.10: Forward rate calculation for money market term

Consider two positions:

- repo of £100 million from 2 January 2000 for 30 days at 6.500%,
- reverse repo of £100 million from 2 January for 60 days at 6.625%.

The two positions can be said to be a 30-day forward 30-day (repo) interest rate exposure (a 30-versus 60-day forward rate). What forward rate must be used if the trader wishes to hedge this exposure, assuming no bid-offer spreads and a 360-day base?

The 30-day by 60-day forward rate can be calculated using the following formula:

$$rf = \left[\left(\frac{1+\left(rs_2 \times \frac{L}{M}\right)}{1+\left(rs_1 \times \frac{S}{M}\right)}\right) - 1\right] \times \frac{M}{L-S}$$

where

rf is the forward rate
rs_2 is the long period rate
rs_1 is the short period rate
L is the long period days
S is the short period days
M is the day-count base

Using this formula we obtain a 30-day by 60-day forward rate of 6.713560%.

Forward rates and compounding

Examples 4.7–4.9 above are for forward rate calculations more than one year into the future, and therefore the formula used must take compounding of interest into consideration. Example 4.10 is for a forward rate within the next 12 months, with one-period bullet interest payments. A different formula is required to account for the sub-one-year periods, as shown in the example.

Case study: Deriving a discount function[25]

In this example we present a traditional bootstrapping technique for deriving a discount function for yield curve fitting purposes. This technique has been called 'naive' (for instance see James and Webber (2000), p. 129) because it suffers from a number of drawbacks. For example it results in an unrealistic forward rate curve, which means that it is unlikely to be used in practice. We review the drawbacks at the end of the case study.

Today is 14 July 2000. The rates given in Table 4.6 are observed in the market. We assume that the day-count basis for the cash instruments and swaps is act/365. Construct the money market discount function.

Creating the discount function

Using the cash money market rates we can create discount factors up to a maturity of six months, using the expression at (4.52):

$$df = \frac{1}{\left(1 + \mathrm{r} \times \frac{days}{365}\right)} \tag{4.54}$$

Table 4.6 Money market rates

Money market rates	*Rate (%)*	*Expiry*	*Days*
1 month	47/32	14/8/00	31
3 months	41/4	16/10/00	94
6 months	41/2	15/1/01	185
Future prices			
Sept 2000	95.60	20/9/00	68
Dec 2000	95.39	20/12/00	169
March 2001	95.25	21/3/01	249
June 2001	94.80	20/6/01	340
Swap rates			
One year (1y)	4.95	16/7/01	367
2 year	5.125	15/7/02	731
3 year	5.28	14/7/03	1095
4 year	5.55	14/7/04	1461
5 year	6.00	14/7/05	1826

[25]In this illustration, the discount function is derived using interest rate data from two off-balance instruments, futures and swaps, as well as money market deposit rates. Derivative instruments are covered in Choudhry (2005).

The resulting discount factors are shown in Table 4.7.

Table 4.7 Discount factors

From	*To*	*Days*	*R%*	*Df*
14 July 2000	14 Aug 2000	31	47/32	0.99642974
	16 Oct 2000	94	41/4	0.98917329
	15 Jan 2001	185	41/2	0.97770040

We can also calculate forward discount factors from the rates implied in the futures prices, which are shown in Table 4.8.

Table 4.8 Forward discount factors

From	*To*	*Days*	*R%*	*Df*
20 Sept 2000	20 Dec 2000	91	4.40	0.98914917
20 Dec 2000	21 March 2001	91	4.61	0.98863717
21 March 2001	20 June 2001	91	4.75	0.98829614
20 June 2001	19 Sept 2000	91	5.20	0.98720154

In order to convert these values into zero-coupon discount factors, we need to first derive a cash 'stub' rate up to the expiry of the first futures contract. The most straightforward way to do this is by linear interpolation of the one-month and three-month rates, as shown in Figure 4.7.

For instance, the calculation for the term marked is

$$4.21875 + \left((4.25 - 4.21875) \times \frac{32}{61} \right) = 4.235143\%$$

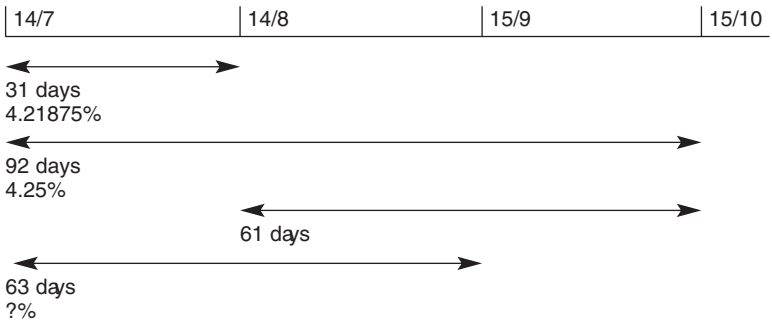


Figure 4.7 Linear interpolation of money and future rates

To convert this to a discount factor:

$$\frac{1}{1+\left(0.04235143\times\frac{63}{365}\right)}=0.99274308$$

From the futures implied forward rates, the zero-coupon discount factors are calculated by successive multiplication of the individual discount factors. These are shown in Table 4.9.

Table 4.9 Zero-coupon discount factors

From	*To*	*Days*	*Df*
14 July 2000	20 Sept 2000	68	0.99172819
	20 Dec 2000	159	0.98172542
	21 March 2001	250	0.96992763
	20 June 2001	341	0.960231459
	19 Sept 2001	432	0.948925494

For the interest-rate swap rates, to calculate discount factors for the relevant dates we use the bootstrapping technique.

One-year swap

We assume a par swap, the present value is known to be 100, and as we know the future value as well, we are able to calculate the one-year zero-coupon rate as shown from the one-year swap rate:

$$df_1=\frac{1}{1+r}=\frac{100}{104.95}=0.95283468$$

Two-year swap

The coupon payment occurring at the end of period one can be discounted back using the one-year discount factor above, leaving a zero-coupon structure as before.

$$df_2=\frac{100-\mathrm{C}\times df_1}{105.125}$$

This gives df_2 equal to 0.91379405.

The same process can be employed for the three, four and five-year par swap rates to calculate the appropriate discount factors.

$$df_3=\frac{100-\mathrm{C}\times(df_1+df_2)}{105.28}$$

This gives df_3 equal to 0.87875624. The discount factors for the four-year and five-year maturities, calculated in the same way, are 0.82899694 and 0.77835621 respectively.

The full discount function is given in Table 4.10 and illustrated graphically in Figure 4.8.

Table 4.10 Discount factors

From	*To*	*Days*	*Zero-coupon (%)*	*Discount factor*	*Source*
14 July 2000	14 Aug 2000	31	4.21875	0.99642974	Money market
	20 Sept 2000	63	4.23500	0.99274308	Money market
	16 Oct 2000	94	4.25000	0.98917329	Money market
	20 Dec 2000	159	4.38000	0.98172542	Futures
	15 Jan 2001	185	4.50000	0.9777004	Money market
	21 March 2001	250	4.55000	0.96992763	Futures
	20 June 2001	341	4.73000	0.96023145	Futures
	16 July 2001	367	4.95000	0.95283468	Swap
	19 Sept 2001	432	5.01000	0.94892549	Futures
	15 July 2002	731	5.12500	0.91379405	Swap
	14 July 2003	1095	5.28000	0.87875624	Swap
	15 July 2004	1461	5.58000	0.82899694	Swap
	15 July 2005	1826	6.10000	0.77835621	Swap

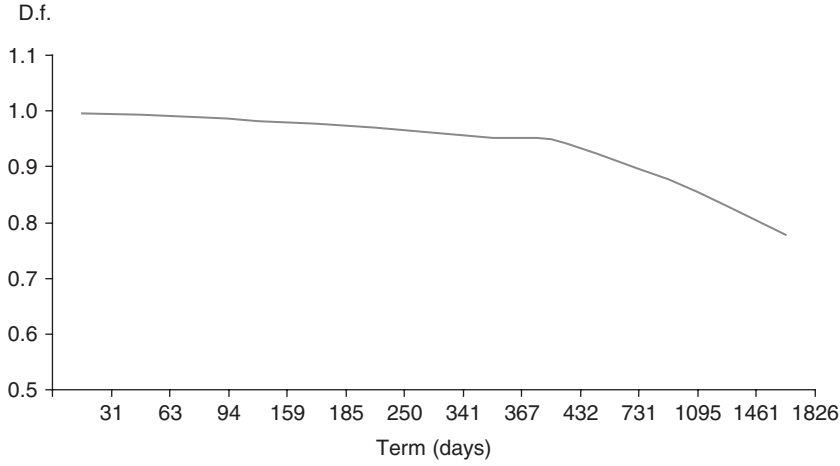


Figure 4.8 Discount equation

Critique of the traditional technique

The method used to derive the discount function in the case study used three different price sources to produce an integrated function and hence yield curve. However there is no effective method by which the three separate curves, which are shown at Figure 4.9, can be integrated into one complete curve. The result is that a curve formed from the three separate curves will exhibit distinct kinks or steps at the points at which one data source is replaced by another data source.

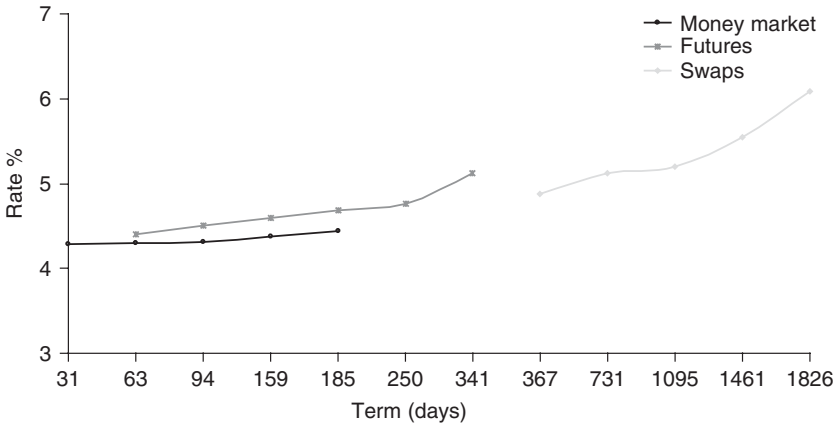


Figure 4.9 Comparison of money market curves

The money market and swap rates incorporate a credit risk premium, reflecting the fact that interbank market counterparties carry an element of default risk. This means that money market rates lie above government repo rates. Futures rates do not reflect default risk, as they are exchange-traded contracts and the exchange clearing house takes on counterparty risk for each transaction. However futures rates are treated as one-point period rates, in effect making them equivalent to forward-rate agreement (FRA) rates. In practice, as the cash flow from FRAs is received as a discounted payoff at one point, whereas futures contract trades require a daily margin payment, a *convexity adjustment* is required to convert futures accurately to FRA rates.

Swap rates also incorporate an element of credit risk, although generally they are considered lower risk as they are off-balance sheet instruments and no principal is at risk. As liquid swap rates are only available for set maturity points, linear interpolation is used to plot points in between available rates. This results in an unstable forward rate curve calculated from the spot rate curve (see James and Webber, 2000), due to the interpolation effect. Nevertheless market makers in certain markets price intermediate-dated swaps based on this linear interpolation method. Another drawback is that the bootstrapping method uses near-maturity rates to build up the curve to far-maturity rates. One of the features of a spot curve derived in this way is that even small changes in short-term rates cause excessive changes in long-dated spot rates, and oscillations in the forward curve. Finally, money market rates beyond the 'stub' period are not considered once the discount factor to the stub date is calculated, so their impact is not felt.

For these reasons the traditional technique, while still encountered in textbooks (including this one) and training courses, is not used very often in the markets.

The theoretical approach described above is neat and appealing, but in practice there are a number of issues that will complicate the attempt to extract zero-coupon rates from bond yields. The main problem is that it is highly unlikely that

we will have a set of bonds that are both precisely six months (or one interest) apart in maturity and priced precisely at par. We also require our procedure to fit as smooth a curve as possible. Setting our coupon bonds at a price of par simplified the analysis in our illustration of bootstrapping, so in reality we need to apply more advanced techniques. A basic approach for extracting zero-coupon bond prices is described in the next section.

Calculating spot rates in practice

Researchers have applied econometric techniques to the problem of extracting a zero-coupon term structure from coupon bond prices. The most well-known approaches are described in McCulloch (1971, 1975), Schaefer (1981), Nelson and Siegel (1987), Deacon and Derry (1994), Adams and Van Deventer (1994) and Waggoner (1997), to name but a few. The most accessible article is probably the one by Deacon and Derry.[26] In addition a good overview of all the main approaches is contained in James and Webber (2000), and chapters 15–18 of their book provide an excellent summary of the research highlights to date.

We have noted that a coupon bond may be regarded as a portfolio of zero-coupon bonds. By treating a set of coupon bonds as a larger set of zero-coupon bonds, we can extract an (implied) zero-coupon interest rate structure from the yields on the coupon bonds.

If the actual term structure is observable, so that we know the prices of zero-coupon bonds of £1 nominal value P_1, P_2, ..., P_N then the price P_C of a coupon bond of nominal value £1 and coupon C is given by

$$P_C = P_1C + P_2C + \ldots + P_n(1 + C) \tag{4.55}$$

Conversely if we can observe the coupon bond yield curve, so that we know the prices P_{C1}, P_{C2},..., P_{CN}, then we may use (4.54) to extract the implied zero-coupon term structure. We begin with the one-period coupon bond, for which the price is

$$P_{C1} = P_1(1 + C)$$

so that

$$P_1 = \frac{P_{C1}}{(1+C)} \tag{4.56}$$

This process is repeated. Once we have the set of zero-coupon bond prices P_1, P_2, ..., P_{N-1} we obtain P_N using

$$P_N = \frac{P_{CN} - P_{N-1}C - \ldots - P_1C}{(1+C)} \tag{4.57}$$

[26]This is in the authors' personal opinion. Those with a good grounding in econometrics will find all these references both readable and accessible. Further recommended references are given in the bibliography.

At this point we apply a regression technique known as *ordinary least squares* (OLS) to fit the term structure. The next chapter discusses this area in greater detail, we have segregated this so that readers who do not require an extensive familiarity with this subject may skip the next chapter. Interested readers should also consult the references at the end of Chapter 5.

Expression (4.55) restricts the prices of coupon bonds to be precise functions of the other coupon bond prices. In fact this is unlikely in practice because specific bonds will be treated differently according to liquidity, tax effects and so on. For this reason we add an error term to (4.55) and estimate the value using cross-sectional regression against all the other bonds in the market. If we say that these bonds are numbered then the regression is given by:

$$P_{C_iN_i} = P_1C_i + P_2C_i + \ldots + P_{N\,i}(1 + C_i) + u_i \tag{4.58}$$

for $i = 1,2,\ldots,I$ and where C_i is the coupon on the ith bond and N_i is the maturity of the ith bond. In (4.58) the regressor parameters are the coupon payments at each interest period date, and the coefficients are the prices of the zero-coupon bonds P_1 to P_N where $j = 1, 2,\ldots,N$. The values are obtained using OLS as long as we have a complete term structure and that $I \geq N$.

In practice we will not have a complete term structure of coupon bonds and so we are not able to identify the coefficients in (4.58). McCulloch (1971, 1975) described a spline estimation method, which assumes that zero-coupon bond prices vary smoothly with term to maturity. In this approach we define P_N, a function of maturity $P(N)$, as a *discount function* given by

$$P(N) = 1 + \sum_{j=1}^{j} a_j f_j(N) \tag{4.59}$$

The function $f_j(N)$ is a known function of maturity N, and the coefficients a_j must be estimated. We arrive at a regression equation by substituting (4.59) into (4.58) to give us (4.60), which can be estimated using OLS.

$$\prod_i = \sum_{j=1}^{J} a_j XL_{ij} + u_{i,} \quad i = 1,2,\ldots,I \tag{4.60}$$

where

$$\prod_i \equiv P_{CiNi} - 1 - C_iN_i$$

$$X_{ij} \equiv f_j(N_i) + C_i \sum_{i=1}^{Ni} f_j(l)$$

The function $f_j(N)$ is usually specified by setting the discount function as a polynomial. In certain texts including McCulloch this is carried out by applying what is known as a *spline* function. Considerable academic research has gone into the use of spline functions as a yield curve fitting technique, however we are not able

to go into the required level of detail here, which is left to the next chapter. Please refer to the bibliography for further information. For a specific discussion on using regression techniques for spline curve fitting methods see Suits *et al*. (1978).

Analysing and interpreting the yield curve

A great deal of effort is expended by bond analysts and economists in analysing and interpreting the shape of the yield curve. This is because the market perceives that there is a considerable information content associated with any yield curve at any time. Here we review the main theories that have been put forward to explain the shape of the yield curve at any one time, all of which have fairly long antecedents. None of the theories can adequately explain everything about yield curves and the shapes they assume at any time, so generally observers seek to explain specific curves using a combination of the accepted theories. This subject is a large one, and it is possible to devote several books to it, so here we seek to introduce the main ideas, with readers directed to the various articles mentioned in the bibliography at the end of the chapter. We assume we are looking at yield curves plotted using risk-free interest rates.

The existence of a yield curve itself indicates that there is a cost associated with funds of different maturities, otherwise we would observe a flat yield curve. The fact that we very rarely observe anything approaching a flat yield curve suggests that investors require different rates of return depending on the maturity of the instrument they are holding.

From observing yield curves in different markets at any time, we notice that a yield curve can adopt one of four basic shapes, which are:

- normal or conventional: in which yields are at 'average' levels and the curve slopes gently upwards as maturity increases, all the way to the longest maturity
- upward-sloping or positive or rising: in which yields are at historically low levels, with long rates substantially greater than short rates
- downward-sloping or inverted or negative: in which yield levels are very high by historical standards, but long-term yields are significantly lower than short rates
- humped: where yields are high with the curve rising to a peak in the medium-term maturity area, and then sloping downwards at longer maturities.

Occasionally yield curves will incorporate a mixture of the above features. For instance a commonly observed curve in developed economies exhibits a positive sloping shape up to the penultimate maturity bond, and then a declining yield for the longest maturity. A diagrammatic representation of each type of curve is given in Figure 4.10.

The expectations hypothesis

Simply put, the *expectations hypothesis* states that the slope of the yield curve reflects the market's expectations about future interest rates. There are in fact four

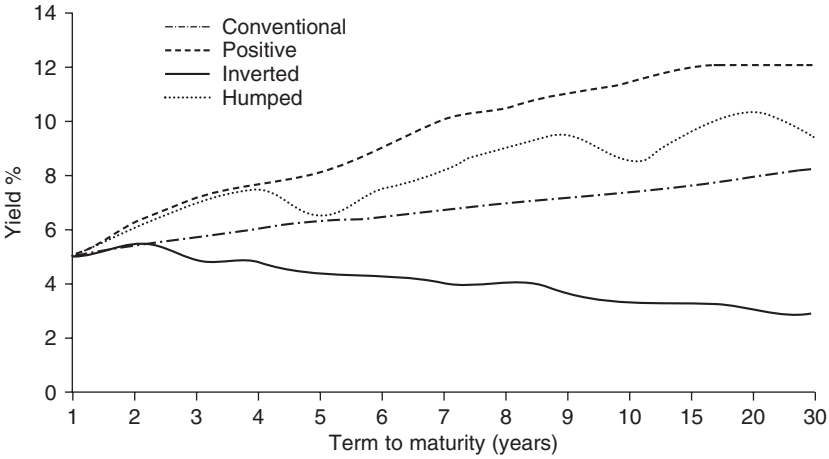


Figure 4.10 The basic shapes of yield curves

main versions of the hypothesis, each distinct from the other – and mutually incompatible.

The expectations hypothesis has a long history, first being described in Fisher (1986) and later developed by Hicks (1946) among others.[27] As Shiller (1990) describes, the thinking behind it probably stems from the way market participants discuss their view on future interest rates when assessing whether to purchase long-dated or short-dated bonds. For instance, if interest rates are expected to fall investors will purchase long-dated bonds in order to 'lock in' the current high long-dated yield. If all investors act in the same way, the yield on long-dated bonds will, of course, decline as prices rise in response to demand; this yield will remain low as long as short-dated rates are expected to fall, and will only revert to a higher level once the demand for long-term rates is reduced. Therefore, downward-sloping yield curves are an indication that interest rates are expected to fall, while an upward-sloping curve reflects market expectations of a rise in short-term interest rates.

The expectations hypothesis suggests that bondholders' expectations determine the course of future interest rates. The two main versions of the hypothesis are the *local expectations hypothesis* and the *unbiased expectations hypothesis*. The *return-to-maturity expectations hypothesis* and *yield-to-maturity expectations hypothesis* are also quoted: for example see Ingersoll (1987).

The unbiased expectations hypothesis states that current forward rates are unbiased predictors of future spot rates. Let $f_1(T,T+1)$ be the forward rate at time t for

[27]See the footnote on page 644 of Shiller (1990) for a fascinating historical note on the origins of the expectations hypothesis. An excellent overview of the hypothesis itself is contained in Ingersoll (1987).

the period from T to $T + 1$. If the one-period spot rate at time T is r_T then according to the unbiased expectations hypothesis:

$$f_1(T,T+1) = E_t[r_T] \tag{4.61}$$

which states that the forward rate $f_1(T,T + 1)$ is the expected value of the future one-period spot rate given by r_T at time T.

The local expectations hypothesis states that all bonds will generate the same expected rate of return if held over a small term. It is given by

$$\frac{E_t[P(t+1,T)]}{P(t,T)} = 1 + r_t \tag{4.62}$$

This version of the hypothesis is the only one that is consistent with no-arbitrage because the expected rates of return on all bonds are equal to the risk-free interest rate. For this reason the local expectations hypothesis is sometimes referred to as the *risk-neutral expectations hypothesis*.

The local expectations hypothesis states that all bonds of the same class, but differing in term to maturity, will have the same expected holding period rate of return. This suggests that a six-month bond and a 20-year bond will produce the same rate of return, on average, over the stated holding period. So if we intend to hold a bond for six months we will receive the same return no matter which specific bond we buy. In general, holding period returns from longer-dated bonds are, on average, higher than those from short-dated bonds. Intuitively we would expect this, with longer-dated bonds offering higher returns to compensate for their higher price volatility (risk). The local expectations hypothesis would not agree with the conventional belief that investors, being risk averse, require higher returns as a reward for taking on higher risk. In addition, it does not provide any insight about the shape of the yield curve. Cox, Ingersoll and Ross (1981) showed that the local expectations hypothesis best reflected equilibrium between spot and forward yields. This was demonstrated using a feature known as Jensen's inequality.

Jarrow (1996, p. 50) states: 'in an economic equilibrium, the returns on ... similar maturity zero-coupon bonds cannot be too different. If they were too different, no investor would hold the bond with the smaller return. This difference could not persist in an economic equilibrium.'

This reflects economic logic, but in practice other factors can impact on holding period returns between bonds that do not have similar maturities. For instance, investors will have restrictions as to which bonds they can hold – banks and building societies are required to hold short-dated bonds for liquidity purposes. In an environment of economic disequilibrium, these investors would still have to hold shorter-dated bonds – even if the holding period return were lower.

So although it is economically neat to expect that the return on a long-dated bond is equivalent to rolling over a series of shorter-dated bonds, it is often observed that longer-term (default-free) returns exceed annualised short-term

default-free returns. So investors who continually rolled over a series of short-dated zero-coupon bonds would most likely receive a lower return than if they had invested in a long-dated zero-coupon bond. Rubinstein (1999) gives an excellent, accessible explanation of why this should be so. The reason is that compared to the theoretical model, future spot rates are not, in reality, known with certainty. This means that short-dated zero-coupon bonds are more attractive to investors for two reasons. First, they are more appropriate instruments to use for hedging purposes; second, they are more liquid instruments, in that they may be more readily converted back into cash than long-dated instruments. With regard to hedging, consider an exposure to rising interest rates. If the yield curve shifts upwards at some point in the future, the price of long-dated bonds will fall by a greater amount. This is a negative result for holders of such bonds, whereas the investor in short-dated bonds will benefit from rolling over his funds at the (new) higher rates. With regard to the second issue, Rubinstein (1999) states:

> It can be shown that in an economy with risk-averse individuals, uncertainty concerning the timing of aggregate consumption, the partial irreversibility of real investments (longer-term physical investments cannot be converted into investments with earlier payouts without sacrifice), [and] ... real assets with shorter-term payouts will tend to have a 'liquidity' advantage.

Therefore the demand for short-term instruments is frequently higher, and hence short-term returns are often lower than long-term returns.

The *pure or unbiased expectations hypothesis* is more commonly encountered, and states that current implied forward rates are unbiased estimators of future spot interest rates.[28] It assumes that investors act in a way that eliminates any advantage of holding instruments of a particular maturity. Therefore if we have a positive-sloping yield curve, the unbiased expectations hypothesis states that the market expects spot interest rates to rise; equally, an inverted yield curve is an indication that spot rates are expected to fall. If short-term interest rates are expected to rise, then longer yields should be higher than shorter ones to reflect this. If this were not the case, investors would only buy the shorter-dated bonds and roll over the investment when they matured. Likewise, if rates are expected to fall then longer yields should be lower than short yields.

The unbiased expectations hypothesis states that the long-term interest rate is a geometric average of expected future short-term rates. This gives us:

$$(1 + rs_N)^N = (1 + rs_1)(1 + {}_1rf_2) \text{ K } (1 + {}_{N-1}rf_N) \tag{4.63}$$

or

$$(1 + rs_N)^N = (1 + rs_{N-1})^{N-1} (1 + {}_{N-1}rf_N) \tag{4.64}$$

[28]For original discussion, see Lutz (1940) and Fisher (1986).

where rs_N is the spot yield on a N-year bond and ${}_{N-1}rf_N$ is the implied one-year rate n years ahead.

For example, if the current one-year spot rate is $rs_1 = 5.0\%$ and the market is expecting the one-year rate in a year's time to be ${}_1rf_2 = 5.539\%$, then the market is expecting a £100 investment in two one-year bonds to yield £100(1.05)(1.05539) = £110.82 after two years. To be equivalent to this an investment in a two-year bond has to yield the same amount, implying that the current two-year rate is $rs_2 = 5.7\%$ as shown below:

$$£100(1 + rs_2)^2 = £110.82$$

which gives us $rs_2 = 5.27\%$, and provides the correct future value as shown below:

$$£100(1.0527)^2 = £110.82$$

This result must be so, to ensure no arbitrage opportunities exist in the market. In fact this is illustrated in elementary texts that discuss and derive forward interest rates. According to the unbiased expectations hypothesis the forward rate ${}_0rf_2$ is an unbiased predictor of the spot rate ${}_1rs_1$ observed one period later; on average the forward rate should equal the subsequent spot rate. The hypothesis can be used to explain any shape in the yield curve.

A rising yield curve is therefore explained by investors expecting short-term interest rates to rise, that is ${}_1rf_2 > rs_2$. A falling yield curve is explained by investors expecting short-term rates to be lower in the future. A humped yield curve is explained by investors expecting short-term interest rates to rise and long-term rates to fall.

Expectations, or views on the future direction of the market, are primarily a function of the expected rate of inflation. If the market expects inflationary pressures in the future, the yield curve will be positively-shaped; if inflation expectations are inclined towards disinflation, then the yield curve will be negative. However, several empirical studies including one by Fama (1976) have shown that forward rates are essentially biased predictors of future spot interest rates – and often overestimate future levels of spot rates.

The unbiased hypothesis has also been criticised for suggesting that investors can forecast (or have a view on) very long-dated spot interest rates, which might be considered slightly unrealistic. As yield curves in most developed-country markets exist to a maturity of up to 30 years or longer, such criticisms may have some substance. Are investors able to forecast interest rates 10, 20 or 30 years into the future? Perhaps not. Nevertheless, this is indeed the information content of, say, a 30-year bond; because the yield on the bond is set by the market, it is valid to suggest that the market has a view on inflation and future interest rates for up to 30 years forward.

The expectations hypothesis is stated in more than one way; other versions include the *return-to-maturity expectations hypothesis*, which states that the total return generated from an investment of term t to T by holding a $(T - t)$-period bond will be equal to the expected return generated by a holding a series of

one-period bonds and continually rolling them over on maturity. More formally we write

$$\frac{1}{P(t,T)} = E_t[(1+r_t)(1+r_{t+1})\ \mathrm{K}\ (1+r_{t-1})] \tag{4.65}$$

The left-hand side of equation (4.65) represents the return received by an investor holding a zero-coupon bond to maturity, which is equal to the expected return associated with rolling over £1 from time t to time T by continually reinvesting one-period maturity bonds, each of which has a yield of the future spot rate r_t.

A related version, the *yield-to-maturity hypothesis*, described in terms of yields, states that the periodic return from holding a zero-coupon bond will be equal to the return from rolling over a series of coupon bonds, but refers to the annualised return earned each year rather than the total return earned over the life of the bond. This assumption enables a zero-coupon yield curve to be derived from the redemption yields of coupon bonds. It is given by

$$\left(\frac{1}{P(t,T)}\right)^{1/(T-t)} = E_t[((1+r_t)(1+r_{t+1})\ \mathrm{K}\ (1+r_{T-1}))^{1/(T-t)}] \tag{4.66}$$

where the left-hand side of equation (4.66) specifies the yield-to-maturity of the zero-coupon bond at time t. In this version the expected holding period *yield* on continually rolling over a series of one-period bonds will be equal to the yield that is guaranteed by holding a long-dated bond until maturity.

The unbiased expectations hypothesis of course states that forward rates are equal to the spot rates expected by the market in the future. The Cox, Ingersoll and Ross (1981) article suggests that only the local expectations hypothesis describes a model that is purely arbitrage-free, as under the other scenarios it would be possible to employ certain investment strategies that would produce returns in excess of what was implied by today's yields. Although it has been suggested[29] that the differences between the local and the unbiased hypotheses are not material, a model that describes such a scenario would not reflect investors' beliefs – which is why further research is ongoing in this area.

The unbiased expectations hypothesis does not, by itself, explain all the shapes of the yield curve or the information content contained within it, so it is often tied in with other explanations, including the liquidity preference theory.

An excellent, accessible overview of all four variants of the expectations hypothesis is given in the Ingersoll (1987) account.

Liquidity preference theory

Intuitively we might feel that longer maturity investments are more risky than shorter ones. An investor lending money for a five-year term will usually demand a higher rate of interest than if he were to lend the same customer money for a

[29] For example, see Campbell (1986); see also Livingstone (1990), pp. 254–6.

five-week term. This is because the borrower might not be able to repay the loan over the longer time period as it might, for instance, have gone bankrupt in that period. For this reason longer-dated yields should be higher than short-dated yields, to recompense the lender for the higher risk exposure during the term of the loan.[30]

We can consider this theory in terms of inflation expectations as well. Where inflation is expected to remain roughly stable over time, the market would anticipate a positive yield curve. However, the expectations hypothesis cannot, by itself, explain this phenomenon – under stable inflationary conditions one would expect a flat yield curve. The risk inherent in longer-dated investments, or the *liquidity preference theory*, seeks to explain a positively-shaped curve.

Generally, borrowers prefer to borrow over as long a term as possible, while lenders will wish to lend over as short a term as possible. Therefore, as we first stated, lenders have to be compensated for lending over the longer term; this compensation is considered a premium for a loss in *liquidity* for the lender. The premium is increased the further the investor lends across the term structure, so that the longest-dated investments will, all else being equal, have the highest yield. So the liquidity preference theory states that the yield curve should almost always be upward-sloping, reflecting bondholders' preference for the liquidity and lower risk of shorter-dated bonds. An *inverted* yield curve could still be explained by the liquidity preference theory when it is combined with the unbiased expectations hypothesis. A *humped* yield curve might be viewed as a combination of an inverted yield curve together with a positive-sloping liquidity preference curve.

The difference between a yield curve explained by unbiased expectations and an actual observed yield curve is sometimes referred to as the *liquidity premium*. This refers to the fact that in some cases short-dated bonds are easier to transact in the market than long-term bonds. It is difficult to quantify the effect of the liquidity premium, which is not static and fluctuates over time. The liquidity premium is so called because, in order to induce investors to hold longer-dated securities, the yields on such securities must be higher than those available on short-dated securities, which are more liquid and may be converted into cash more easily. The liquidity premium is the compensation required for holding less liquid instruments. If longer-dated securities then provide higher yields, as is suggested by the existence of the liquidity premium, they should generate, on average, higher total returns over an investment period. This is not consistent with the local expectations hypothesis. More formally we can write:

$$0 = L_1 < L_2 < L_3 < \ L < L_n \text{ and } (L_2 - L_1) > (L_3 - L_2) > L\ (L_n - \ L_{n-1})$$

where L is the premium for a bond with term to maturity of n years; this states that the premium increases as the term to maturity rises and that an otherwise flat yield curve will have a positively-sloping curve, with the degree of slope steadily decreasing as we extend along the yield curve. This is consistent with observation of yield curves under 'normal' conditions.

[30]For original discussion, see Hicks (1946).

The expectations hypothesis assumes that forward rates are equal to the expected future spot rates, as shown by equation (4.67):

$$_{n-1}rf_n = E(_{n-1}rs_n) \tag{4.67}$$

where $E(\ )$ is the expectations operator for the current period. This assumption implies that the forward rate is an unbiased predictor of the future spot rate, as we suggested in the previous paragraph. Liquidity preference theory, on the other hand, recognises the possibility that the forward rate may contain an element of liquidity premium which declines over time as the starting period approaches, given by equation (4.68):

$$_{n-1}rf_n > E(_{n-1}rs_n) \tag{4.68}$$

If there was uncertainty in the market about the future direction of spot rates – and hence where the forward rate should lie – equation (4.67) is adjusted to give the reverse inequality.

Figure 4.11 illustrates both expectations hypothesis and liquidity preference impact on the yield curve.

Money substitute hypothesis

A particular explanation of short-dated bond yield curves has been attempted by Kessel (1965). In the *money substitute theory* short-dated bonds are regarded as substitutes for holding cash. Investors hold only short-dated market instruments because these are viewed as low or negligible risk. As a result the yields of short-dated bonds are depressed due to the increased demand and lie below longer-dated bonds. Borrowers, on the other hand, prefer to issue debt for longer maturities, and on as few occasions as possible, to minimise funding costs and reduce uncertainty. Therefore, the yields of longer-dated paper are driven upwards due to a combination of increased supply and lower liquidity. In certain respects

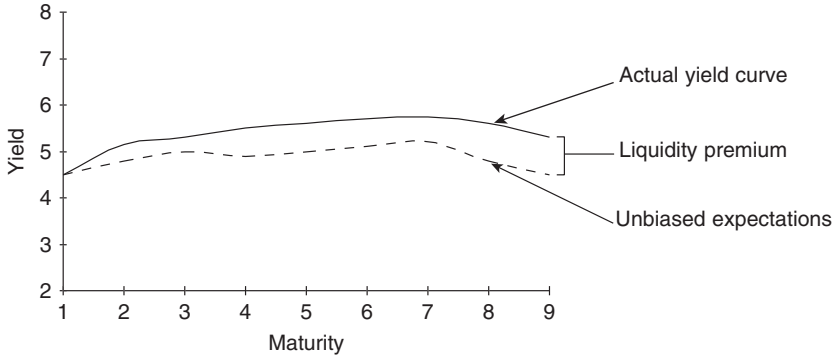


Figure 4.11 Yield curve explained by expectations hypothesis and liquidity preference

the money substitute theory is closely related to the liquidity preference theory, and by itself does not explain inverted or humped yield curves.

Segmentation hypothesis

The capital markets are made up of a wide variety of users, each with different requirements. Certain classes of investors will prefer dealing at the short end of the yield curve, while others will concentrate on the longer end of the market. The *segmented markets theory* suggests that activity is concentrated in certain specific areas of the market, and that there are no interrelationships between these parts of the market; the relative amounts of funds invested in each area of the maturity spectrum causes differentials in supply and demand, which results in humps in the yield curve. That is, the shape of the yield curve is determined by supply and demand for certain specific maturity investments, each of which has no reference to any other part of the curve.

The segmented markets hypothesis seeks to explain the shape of the yield curve by stating that different types of market participants invest in different sectors of the term structure, according to their requirements. So, for instance, the banking sector has a requirement for short-dated bonds, while pension funds will invest in the long-end of the market. This was first described in Culbertson (1957).

There may also be regulatory reasons for different investors to have preferences for particular maturity investments. So, for example, banks and building societies concentrate a large part of their activity at the short end of the curve, as part of daily cash management (known as *asset and liability management*) and for regulatory purposes (known as *liquidity requirements*). Fund managers such as pension funds and insurance companies are active at the long end of the market. Few institutional investors, however, have any preference for medium-dated bonds. This behaviour on the part of investors will lead to high prices (low yields) at both the short and long ends of the yield curve and lower prices (higher yields) in the middle of the term structure.

According to the segmented markets hypothesis, a separate market exists for specific maturities along the term structure, thus interest rates for these maturities are set by supply and demand (see Culbertson 1957). Where there is no demand for a particular maturity, the yield will lie above other segments. Market participants do not hold bonds in any other area of the curve outside their area of interest[31] so that short-dated and long-dated bond yields exist independently of each other.

The segmented markets theory is usually illustrated by reference to banks and life companies. Banks and building societies hold their funds in short-dated instruments, usually no longer than five years in maturity. This is because of the nature of retail banking operations, with a large volume of instant access funds being deposited at banks, and also for regulatory purposes. Holding short-term, liquid bonds enables banks to meet any sudden or unexpected demand for funds from customers. The classic theory suggests that as banks invest their funds in

[31] For example, retail and commercial banks hold bonds in the short dates, while life assurance companies hold long-dated bonds.

short-dated bonds, the yields on these bonds is driven down. When they subsequently liquidate part of their holding, perhaps to meet higher demand for loans, the yields are driven up and prices of the bonds fall. This affects the short end of the yield curve but not the long end.

The segmented markets theory can be used to cover an explanation of any particular shape of the yield curve, although it may be argued that it fits best with positive-sloping curves. However, it does not offer us any help to *interpret* the yield curve whatever shape it may be, and therefore offers no information content during analysis. By definition the theory suggests that for investors, bonds with different maturities are not perfect substitutes for each other. This is because different bonds would have different holding period returns, making them imperfect substitutes of one another. As a result of bonds being imperfect substitutes, markets are segmented according to maturity.

The segmentations hypothesis is a reasonable explanation of certain features of a conventional positively-sloping yield curve, but by itself is not sufficient. There is no doubt that banks and building societies have a requirement to hold securities at the short end of the yield curve, as much for regulatory purposes as for yield considerations; however, other investors are probably more flexible and will place funds where value is deemed to exist. Nevertheless, the higher demand for benchmark securities does drive down yields along certain segments of the curve.

A slightly modified version of the market segmentation hypothesis is known as the *preferred habitat theory*, first described in Modigliani and Sutch (1966), which states not only that investors have a preferred maturity but also that they may move outside this sector if they receive a premium for so doing. This would explain 'humped' shapes in yield curves.

This suggests that different market participants have an interest in specified areas of the yield curve, but can be induced to hold bonds from other parts of the maturity spectrum if there is sufficient incentive. Hence banks may, at certain times, hold longer-dated bonds once the price of these bonds falls to a certain level, making the return on the bonds worth the risk involved in holding them. Similar considerations may persuade long-term investors to hold short-dated debt. So higher yields will be required to make bondholders shift out of their usual area of interest. This theory essentially recognises the flexibility that investors have, outside regulatory or legal requirements (such as the terms of an institutional fund's objectives), to invest in whatever part of the yield curve they identify as offering value. The preferred habitat theory may be viewed as a version of the liquidity preference hypothesis, where the preferred habitat is the short end of the yield curve, so that longer-dated bonds must offer a premium in order to entice investors to hold them. This is described in Cox, Ingersoll and Ross (1981).

The combined theory

The explanation for the shape of the yield curve at any time is more likely to be given by a combination of the pure expectations hypothesis and the liquidity preference theory, and possibly one or two other theories. Market analysts often combine the unbiased expectations hypothesis with the liquidity preference

theory into an 'eclectic' theory. The result is fairly consistent with any shape of yield curve, and is also a predictor of rising interest rates. In the combined theory the forward interest rate is equal to the expected future spot rate, together with a quantified liquidity premium. This is shown by equation (4.69):

$$_0rf_i = E(_{i-1}rs_1) = L_i \tag{4.69}$$

where L_i is the liquidity premium for a term to maturity of i years. The size of the liquidity premium is expected to increase with increasing maturity, so that $L_i > L_{i-1}$. An example is given below.

Consider the interest rate structure in Table 4.11. The current term structure is positive-sloping since the spot rates increase with increasing maturity. However, the market expects future spot rates to be constant at 4.5%. The forward and spot rates are also shown; however, the forward rate is a function of the expected spot rate and the liquidity premium. This premium is equal to 0.50% for the first year, 1.0% in the second and so on.

Table 4.11 Positive yield curve with constant expected future rates (%)

Period n	0	1	2	3	4	5
$E(rs)$		4.5	4.5	4.5	4.5	4.5
Forward rate $_0rf_n$		5.0	5.5	6.0	6.5	7.5
Spot rate rs_n	5.0	5.3	5.8	6.2	6.8	7.0

The combined theory is consistent with an inverted yield curve. This will apply even when the liquidity premium is increasing with maturity; for example, where the expected future spot interest rate is declining. Typically this would be where there was a current term structure of falling yields along the term structure. The spot rates might be declining where the fall in the expected future spot rate exceeds the corresponding increase in the liquidity premium.

The flat yield curve

The conventional theories do not seek to explain a flat yield curve. Although it is rare to observe flat curves in a market, certainly for any length of time, they do emerge occasionally in response to peculiar economic circumstances. Conventional thinking contends that a flat curve is not tenable because investors should, in theory, have no incentive to hold long-dated bonds over shorter-dated bonds when there is no yield premium, so that as they sell off long-dated paper the yield at the long-end should rise, producing an upward-sloping curve. In previous circumstances of a flat curve, analysts have produced different explanations for their existence. In November 1988 the US Treasury yield curve was flat relative to the recent past; researchers contended that this was the result of the market's view that long-dated yields would fall as bond prices rallied upwards.[32]

[32]See Levy (1999).

One recommendation is to buy longer maturities when the yield curve is flat, in anticipation of lower long-term interest rates, which is the direct opposite to the view that a flat curve is a signal to sell long bonds. In the case of the US market in 1988, long bond yields did in fact fall by approximately 2% in the following 12 months. This would seem to indicate that one's view of future long-term rates should be behind the decision to buy or sell long bonds, rather than the shape of the yield curve itself. A flat curve may well be more heavily influenced by supply and demand factors than anything else, with the majority opinion eventually winning out and forcing a change in the curve to a more conventional shape.

Further views on the yield curve

In this discussion we assume an economist's world of the *perfect market* (also sometimes called the *frictionless* financial market). Such a perfect capital market is characterised by:

- perfect information
- no taxes
- bullet maturity bonds
- no transaction costs.

Of course, in practice markets are not completely perfect. However, assuming perfect markets makes the discussion of spot and forward rates and the term structure easier to handle. When we analyse yield curves for their information content, we have to remember that the markets which they represent are not perfect, and that frequently we observe anomalies that are not explained by the conventional theories.

At any one time it is probably more realistic to suggest that a range of factors contributes to the yield curve being one particular shape. For instance, short-term interest rates are greatly influenced by the availability of funds in the money market. The slope of the yield curve (usually defined as the ten-year yield minus the three-month interest rate) is also a measure of the degree of tightness of government monetary policy. A low, upward-sloping curve is often thought to be a sign that an environment of cheap money, due to a looser monetary policy, is to be followed by a period of higher inflation and higher bond yields. Equally, a high downward-sloping curve is taken to mean that a situation of tight credit, due to a stricter monetary policy, will result in falling inflation and lower bond yields. Inverted yield curves have often preceded recessions; for instance, the *Economist* in an article from April 1998 remarked that, with one exception, every recession in the United States since 1955 had been preceded by a negative yield curve.[33] The analysis is the same: if investors expect a recession they also expect inflation to fall, so the yields on long-term bonds will fall relative to short-term bonds. So the conventional explanation for an inverted yield curve is that the markets and the

[33]The exception was the one precipitated by the 1973 oil shock.

investment community expect either a slow-down of the economy, or an outright recession.[34] In this case one would expect the monetary authorities to ease the money supply by reducing the base interest rate in the near future: hence an inverted curve. At the same time, a reduction of short-term interest rates will affect short-dated bonds and these are sold off by investors, further raising their yield.

While the conventional explanation for negative yield curves is an expectation of economic slow-down, on occasion other factors will be involved. In the UK during the period from July 1997 to June 1999 the gilt yield curve was inverted.[35] There was no general view that the economy was heading for recession; in fact, the newly elected Labour government inherited an economy believed to be in satisfactory shape. Instead, the explanation behind the inverted shape of the gilt yield curve focused on two other factors: first, the handing of responsibility for setting interest rates to the Monetary Policy Committee (MPC) of the Bank of England; secondly, the expectation that the UK would, over the medium term, abandon sterling and join the euro currency. The yield curve at this time suggested that the market expected the MPC to be successful and keep inflation at a level of around 2.5% over the long term (its target is actually a 1% range either side of 2.5%), and also that sterling interest rates would need to come down over the medium term as part of *convergence* with interest rates in Euroland. These are both medium-term expectations however, and, in the authors' view, are not logical at the short end of the yield curve. In fact the term structure moved to a positive-sloped shape up to the six to seven-year area, before inverting out to the long end of the curve, in June 1999. This is a more logical shape for the curve to assume, but it was short-lived and returned to being inverted after the two-year term.

There is, therefore, significant information content in the yield curve, and economists and bond analysts will consider the shape of the curve as part of their policy-making and investment advice. The shape of parts of the curve, whether the short end or long end, as well that of the entire curve, can serve as useful predictors of future market conditions. As part of an analysis it is also worthwhile considering the yield curves across several different markets and currencies. For instance, the interest-rate swap curve, and its position relative to that of the government bond yield curve, are also regularly analysed for their information content. In developed-country economies the swap market is invariably as liquid as the government bond market, if not more liquid, and so it is common to see

[34] A recession is formally defined as two successive quarters of falling output in the domestic economy.

[35] Although the gilt yield curve changed to being positively-sloped out to the seven to eight-year maturity area, for a brief period in June–July 1999, it very quickly reverted to being inverted throughout the term structure, and remained so until May/June 2001, when it changed once again to being slightly positive-sloping up to the four-year term, and inverting from that point onwards. This shape at least is more logical and explainable (and had been called by the author for the preceding two years!).

the swap curve analysed when making predictions about, say, the future level of short-term interest rates.

Government policy will influence the shape and level of the yield curve, including policy on public sector borrowing, debt management and open-market operations.[36] The market's perception of the size of public sector debt will influence bond yields; for instance, an increase in the level of debt can lead to an increase in bond yields across the maturity range. Open-market operations can have a number of effects. In the short term it can tilt the yield curve both upwards and downwards; longer term, changes in the level of the base rate will affect yield levels. An anticipated rise in base rates can lead to a drop in prices for short-term bonds, whose yields will be expected to rise; this can lead to a (temporary) inverted curve. Finally, debt management policy[37] will influence the yield curve. Much government debt is rolled over as it matures, but the maturity of the replacement debt can have a significant influence on the yield curve in the form of humps in the market segment in which the debt is placed, if the debt is priced by the market at a relatively low price and hence high yield.

Selected bibliography and references

Adams, K. and Van Deventer, D. 'Fitting yield curves and forward rate curves with maximum smoothness', *Journal of Fixed Income*, 4, 1994, pp. 52–62.

Avellaneda, M. *Quantitative Modelling of Derivative Securities*, Chapman & Hall/CRC, 2000, chs 10–12.

Baxter, M. and Rennie, A. *Financial Calculus*, Cambridge University Press, 1996, ch. 5.

Campbell J. 'A defence of traditional hypotheses about the term structure of interest rates,' *Journal of Finance*, March 1986, pp. 183–93.

Campbell, J., Lo, A. and McKinlay, A. *The Econometrics of Financial Markets*, Princeton University Press, 1997, chs 10–11.

Choudhry, M. *Bond Market Securities*, FT Prentice Hall, 2001, chs 4 and 5.

Choudhry, M. *Fixed Income Markets: Instruments, Applications, Mathematics*, Singapore: Wiley, 2005.

Cox, J., Ingersoll, J. and Ross, S. 'An inter-temporal general equilibrium model of asset prices', *Econometrica*, 53, 1985.

Culbertson, J. M. 'The term structure of interest rates,' *Quarterly Journal of Economics*, Nov. 1957, pp. 485–517.

Deacon, M. and Derry, A., 'Estimating the term structure of interest rates', Bank of England Working Paper Series, No. 24, July 1994.

Duffie, D. *Dynamic Asset Pricing Theory*, Princeton University Press, 1992.

Eales, B. and Choudhry, M. *Derivative Instruments*, Butterworth Heinemann, 2003.

Fabozzi, F. *Bond Markets, Analysis and Strategies*, 2nd edn, Prentice Hall, 1993, ch. 5.

Fabozzi, F. *Fixed Income Mathematics*, McGraw-Hill, 1997.

Fabozzi, F. (ed.) *The Handbook of Fixed Income Securities*, 5th edn, McGraw-Hill, 1997.

Fabozzi, F. *Valuation of Fixed Income Securities and Derivatives*, 3rd edn, FJF Associates, 1998.

Fabozzi, F. *Treasury Securities and Derivatives*, FJF Associates, 1998.

[36] 'Open-market operations' refers to the daily operation by the Bank of England to control the level of the money supply (to which end the Bank purchases short-term bills and also engages in repo dealing).

[37] In the United Kingdom this is now the responsibility of the Debt Management Office.

Fama, E.F. 'The information in the term structure,' *Journal of Financial Economics*, Dec. 1984, pp. 509–28.
Fisher I. 'Appreciation of interest,' Publications of the American Economic Association, August 1986, pp. 23–39.
Guajarati, D. *Basic Econometrics*, 3rd edn, McGraw Hill, 1995.
Hicks, J. *Value and Capital*, 2nd edn, Oxford University Press, 1946.
Ingersoll, J. Jr., *Theory of Financial Decision Making*, Rowman & Littlefield, 1987, ch. 18.
Ingersoll, J., *Theory of Financial Decision Making*, Bowman & Littlefield 1987, Chapter 18.
James, J. and Webber, N. *Interest Rate Modelling*, Wiley, 2000.
Jarrow, R. *Modelling Fixed Income Securities and Interest Rate Options*, McGraw-Hill, 1996, ch. 3.
Jarrow, R., *Modelling Fixed Income Securities and Interest Rate Options*, McGraw-Hill 1996, Chapters 2–3.
Kessel, R.A. 'The cyclical behaviour of the term structure of interest rates,' in *Essays in Applied Price Theory*, University of Chicago, 1965.
Kitter, G. *Investment Mathematics for Finance and Treasury Professionals*, Wiley, 1999, chs 3 and 5.
Levy H. *Introduction to Investments*, 2nd edn, South-Western College Publishing, 1999.
Livingstone, M. *Money and Capital Markets*, Prentice-Hall, 1990.
Lutz, F. 'The structure of interest rates,' *Quarterly Journal of Economics*, Nov. 1940, pp. 36–63.
Martellini, L., Priaulet, S. and Priaulet, P. *Fixed-Income Securities: Valuation, Risk Management and Portfolio Strategies,* Chichester: Wiley, 2003.
McCulloch, J. 'Measuring the term structure of interest rates', *Journal of Business*, 44, 1971, pp. 19–31.
McCulloch, J. 'The tax-adjusted yield curve', *Journal of Finance*, 30, 1975, pp. 811–30.
Modigliani, F. and Sutch, R. 'Innovations in interest rate policy,' *American Economic Review*, Vol. 56, 1966, pp. 178–97.
Neftci, S. *An Introduction to the Mathematics of Financial Derivatives*, 2nd edn, Academic Press, 2000, ch. 18.
Nelson, C. and Siegel, A. 'Parsimonious modelling of yield curves', *Journal of Business*, Vol. 60, No. 4, 1987, pp. 473–89.
Questa, G. *Fixed Income Analysis for the Global Financial Market*, Wiley, 1999.
Ross, S. M. *An Introduction to Mathematical Finance*, Cambridge University Press, 1999.
Rubinstein, M. 'Rubinstein on derivatives,' *Risk,* 1999, pp. 84–5.
Schaefer, S. 'Measuring a tax-specific term structure of interest rates in the market for British government securities', *Economic Journal*, 91, 1981, pp. 415–38.
Shiller, R. 'The term structure of interest rates', in Friedman, B. and Hanh, F. (eds), *Handbook of Monetary Economics*, Vol. 1, North Holland, 1990, ch. 13.
Suits, D., Mason, A. and Chan, L. 'Spline functions fitted by standard regression methods', *Review of Economics and Statistics*, 60, 1978, pp. 132–9.
Sundaresan, S. *Fixed Income Markets and Their Derivatives*, South-Western, 1997.
Tuckman, B. *Fixed Income Securities*, Wiley, 1996.
Van Deventer, D. and Imai, K. *Financial Risk Analytics*, Irwin, 1997, pp. 9–11.
Waggoner, D. 'Spline methods for extracting interest rate curves from coupon bond prices', Working Paper, Federal Reserve Bank of Atlanta, 97–10, 1997.
Windas, T. *An Introduction to Option-Adjusted Spread Analysis*, Bloomberg, 1993, ch. 5.

In addition interested readers may wish to consult the following recommended references on term structure analysis.

Constantinides, G. 'A theory of the nominal term structure of interest rates', *Review of Financial Studies*, Vol. 5, No. 4, 1992, pp. 531–52.
Cox, J., Ingersoll, J. and Ross, S. 'A re-examination of traditional hypotheses about the term structure of interest rates', *Journal of Finance*, 36, 1981, pp. 769–99.

Cox, J., Ingersoll, J. and Ross, S., 'A theory of the term structure of interest rates', *Econometrica*, 53, 1985, pp. 385–407.
Culbertson, J. 'The term structure of interest rates', *Quarterly Journal of Economics,* 71, 1957, pp. 489–504.
McCulloch, J.H. 'A reexamination of traditional hypotheses about the term structure: a comment', *Journal of Finance*, Vol. 63, No. 2, 1993, pp. 779–89.
Shiller, R., Campbell, J. and Schoenholtz, K. 'Forward rates and future policy: interpreting the term structure of interest rates', *Brookings Papers on Economic Activity*, 1, 1983, pp. 173–223.
Stambaugh, R. 'The information in forward rates: implications for models of the term structure', *Journal of Financial Economics*, 21, 1988, pp. 41–70.

5
Fixed Income Securities II: Interest-Rate Risk

In this chapter we discuss the sensitivity of bond prices to changes in market interest rates, and the key concepts of duration and convexity.

Duration, modified duration and convexity

Bonds pay a part of their total return during their lifetime, in the form of coupon interest, so that the term to maturity does not reflect the true period over which the bond's total return is earned. Additionally if we wish to gain an idea of the trading characteristics of a bond, and compare this with other bonds of say, similar maturity, term to maturity is insufficient and so we need a more accurate measure. A plain vanilla coupon bond pays out a proportion of its return during the course of its life, in the form of coupon interest. If we were to analyse the properties of a bond, we should conclude quite quickly that its maturity does not give us much indication of how much of its return is paid out during its life, or any idea of the timing or size of its cash flows, and hence its sensitivity to moves in market interest rates. For example, we might compare two bonds with the same maturity date but different coupons, where the higher-coupon bond provides a larger proportion of its return in the form of coupon income than does the lower-coupon bond. The higher-coupon bond provides its return at a faster rate; its value is theoretically therefore less subject to subsequent fluctuations in interest rates.

We may wish to calculate an average of the time to receipt of a bond's cash flows, and use this measure as a more realistic indication of maturity. However cash flows during the life of a bond are not all equal in value, so a more accurate measure would be to take the average time to receipt of a bond's cash flows, but weighted in the form of the cash flows' present value. This is, in effect, *duration*. We can measure the speed of payment of a bond, and hence its price risk relative to other bonds of the same maturity, by measuring the average maturity of the bond's cash flow stream. Bond analysts use duration to measure this property (it is sometimes known as *Macaulay's duration*, after its inventor, who first introduced it in 1938: see Macaulay (1938)). Duration is the weighted average time until the receipt of cash flows from a bond, where the weights are the present values of the cash flows, measured in years. At the time that he introduced the concept,

Macaulay used the duration measure as an alternative for the length of time that a bond investment had remaining to maturity.

Duration

The price/yield formula for a plain vanilla bond is as given at (5.1), assuming complete years to maturity paying annual coupons, and with no accrued interest at the calculation date. Note that the symbol for yield to maturity here is *r*.

$$\mathrm{P} = \frac{C}{(1+r)} + \frac{C}{(1+r)^2} + \frac{C}{(1+r)^3} + \ldots + \frac{C}{(1+r)^n} + \frac{M}{(1+r)^n} \tag{5.1}$$

If we take the first derivative of this expression we obtain (5.2):

$$\frac{dP}{dr} = \frac{(-1)C}{(1+r)^2} + \frac{(-2)C}{(1+r)^3} + \ldots + \frac{(-n)C}{(1+r)^{n+1}} + \frac{(-n)M}{(1+r)^{n+1}} \tag{5.2}$$

If we rearrange (5.2) we obtain the expression at (5.3), which is the equation to calculate the approximate change in price for a small change in yield:

$$\frac{dP}{dr} = \frac{1}{(1+r)}\left[\frac{1C}{(1+r)} + \frac{2C}{(1+r)^2} + \ldots + \frac{nC}{(1+r)^n} + \frac{nM}{(1+r)^n}\right] \tag{5.3}$$

Readers may feel a sense of familiarity regarding the expression in brackets in equation (5.3) as this is the weighted average time to maturity of the cash flows from a bond, where the weights are, as in our example above, the present values of each cash flow. The expression at (5.3) gives us the approximate measure of the change in price for a small change in yield. If we divide both sides of (5.3) by *P* we obtain the expression for the approximate percentage price change, given at (5.4).

$$\frac{dP}{dR}\frac{1}{P} = -\frac{1}{(1+r)}\left[\frac{1C}{(1+r)} + \frac{2C}{(1+r)^2} + \ldots + \frac{nC}{(1+r)^n} + \frac{nM}{(1+r)^n}\right]\frac{1}{P} \tag{5.4}$$

If we divide the bracketed expression in (5.4) by the current price of the bond *P* we obtain the definition of Macaulay duration, given at (5.5):

$$D = \frac{\dfrac{1C}{(1+r)} + \dfrac{2C}{(1+r)^2} + \ldots + \dfrac{nC}{(1+r)^n} + \dfrac{nM}{(1+r)^n}}{P} \tag{5.5}$$

Equation (5.5) is simplified using $\sum$ as shown by (5.6):

$$D = \frac{\displaystyle\sum_{n=1}^{N} \frac{nC_n}{(1+r)^n}}{P} \tag{5.6}$$

where *C* represents the bond cash flow at time *n*.

The Macaulay duration value given by (5.6) is measured in years. However as a measure of interest-rate sensitivity, and for use in hedge calculations, duration can

be transformed into *modified duration*. This was the primary measure of interest rate risk used in the markets, and is still widely used despite the advent of the *value-at-risk* measure for market risk.

If we substitute the expression for Macaulay duration (5.5) into equation (5.4) for the approximate percentage change in price we obtain (5.7):

$$\frac{dP}{dr}\frac{1}{P} = -\frac{1}{(1+r)}D \tag{5.7}$$

This is the definition of modified duration, given as (5.8):

$$MD = \frac{D}{(1+r)} \tag{5.8}$$

Modified duration is clearly related to duration, and we can use it to indicate that, for small changes in yield, a given change in yield results in an inverse change in bond price. We can illustrate this by substituting (5.8) into (5.7), giving us (5.9):

$$\frac{dP}{dr}\frac{1}{P} = -MD \tag{5.9}$$

If we are determining duration long-hand, there is another arrangement we can use to shorten the procedure. Instead of equation (5.1) we use (5.10) as the bond price formula, which calculates price based on a bond as being comprised of an annuity stream and a redemption payment, and sums the present values of these two elements. Again we assume an annual coupon bond priced on a date that leaves a complete number of years to maturity and with no interest accrued.

$$P = C\left[\frac{1-\frac{1}{(1+r)^n}}{r}\right] + \frac{M}{(1+r)^n} \tag{5.10}$$

This expression calculates the price of a bond as the present value of the stream of coupon payments and the present value of the redemption payment. If we take the first derivative of (5.10) and then divide this by the current price of the bond P, the result is another expression for the modified duration formula, given at (5.11).

$$MD = \frac{\frac{C}{r^2}\left[1-\frac{1}{(1+r)^n}\right] + \frac{n\left(M-\frac{C}{r}\right)}{(1+r)^{n+1}}}{P} \tag{5.11}$$

For an irredeemable bond duration is given by:

$$D = \frac{1}{rc} \tag{5.12}$$

where $rc = (C/P_d)$ is the *running yield* (or *current yield*) of the bond. This follows from equation (5.6) as $N \to \infty$, recognising that for an irredeemable bond $r = rc$ (rc is the

bond's current or flat yield). Equation (5.12) provides the limiting value to duration. For bonds trading at or above par, duration increases with maturity and approaches this limit from below. For bonds trading at a discount to par, duration increases to a maximum at around 20 years and then declines towards the limit given by (5.12). So in general duration increases with maturity, with an upper bound given by (5.12).

Properties of Macaulay duration

A bond's duration is always less than its maturity. This is because some weight is given to the cash flows in the early years of the bond's life, and this brings forward the average time at which cash flows are received. In the case of a zero-coupon bond, there is no present value weighting of the cash flows, for the simple reason that there are no cash flows, and so duration for a zero-coupon bond is equal to its term to maturity. Duration varies with coupon, yield and maturity. The following three factors imply higher duration for a bond:

- the lower the coupon
- the lower the yield
- broadly, the longer the maturity.

Duration increases as coupon and yield decrease. As the coupon falls, more of the relative weight of the cash flows is transferred to the maturity date, and this causes duration to rise. Because the coupon on index-linked bonds is generally much lower than on vanilla bonds, this means that the duration of index-linked bonds will be much higher than for vanilla bonds of the same maturity. As yield increases, the present values of all future cash flows fall, but the present values of the more distant cash flows fall relatively more than those of the nearer cash flows. This has the effect of increasing the relative weight given to nearer cash flows and hence of reducing duration.

Modified duration

Although it is common for newcomers to the market to think intuitively of duration much as Macaulay originally did, as a proxy measure for the time to maturity of a bond, such an interpretation misses the main point of duration, which is a measure of price volatility or interest rate risk.

Using the first term of a Taylor's expansion of the bond price function[1] we can show the relationship between price volatility and the duration measure expressed in (5.13):

$$\Delta P = -\left[\frac{1}{(1+r)}\right] \times \text{Macaulay duration} \times \text{Change in yield} \tag{5.13}$$

where r is the yield to maturity for an annual-paying bond (for a semi-annual coupon bond, we use $\frac{r}{2}$). If we combine the first two components of the right-hand

[1] For an accessible explanation of the Taylor expansion, see Butler (1998), pp. 112–14.

side, we obtain the definition of modified duration. Equation (5.13) expresses the approximate percentage change in price as being equal to the modified duration multiplied by the change in yield. We saw in the previous section how the formula for Macaulay duration could be modified to obtain the *modified duration* for a bond. There is a clear relationship between the two measures. From the Macaulay duration of a bond can be derived its modified duration, which gives a measure of the sensitivity of a bond's price to small changes in yield. As we have seen, the relationship between modified duration and duration is given by (5.14).

$$MD = \frac{D}{(1+r)} \tag{5.14}$$

where *MD* is the modified duration in years. However it also measures the approximate change in bond price for a 1% change in bond yield. For a bond that pays semi-annual coupons, the equation becomes:

$$MD = \frac{D}{(1+\frac{1}{2}r)} \tag{5.15}$$

This means that the following relationship holds between modified duration and bond prices:

$$\Delta P = MD \times \Delta r \times P \tag{5.16}$$

In the UK markets the term *volatility* is sometimes used to refer to modified duration, but this is becoming increasingly uncommon in order to avoid confusion with option markets' use of the same term, which there refers to *implied volatility* and is something quite different.

Example 5.1: Using modified duration

An 8% annual coupon bond is trading at par with a duration of 2.85 years. If yields rise from 8% to 8.50%, then the price of the bond will fall by:

$$\begin{aligned}\Delta P &= -\mathrm{D} \times \frac{\Delta r}{(1+r)} \times P \\ &= -(2.85) \times \left(\frac{0.005}{1.080}\right) \times 100 \\ &= -£1.3194\end{aligned}$$

That is, the price of the bond will now be £98.6806.

The modified duration of a bond with a duration of 2.85 years and yield of 8% is obviously:

$$\mathrm{MD} = \frac{2.85}{1.08}$$

which gives us *MD* equal to 2.639 years.

Consider a five-year 8% annual bond priced at par with a duration of 4.31 years. The modified duration can be calculated to be 3.99. This tells us that for a 1% move in the yield to maturity, the price of the bond will move (in the opposite direction) by 3.99%.

We can use modified duration to approximate bond prices for a given yield change. This is illustrated with the following expression:

$$\Delta P = -MD \times (\Delta r) \times P \tag{5.17}$$

For a bond with a modified duration of 3.99, priced at par, an increase in yield of one basis point (100 basis = 1%) leads to a fall in the bond's price of:

$$\Delta P = (-3.24/100) \times (+0.01) \times 100.00$$
$$\Delta P = £0.0399, \text{ or } 3.99 \text{ pence.}$$

In this case 3.99 pence is the *basis point value (BPV)* of the bond, which is the change in the bond price given a one basis point change in the bond's yield. The basis point value of a bond can be calculated using (5.18).

$$BPV = \frac{MD}{100} \bullet \frac{P}{100} \tag{5.18}$$

BPVs are used in hedging bond positions. To hedge a bond position requires an opposite position to be taken in the hedging instrument, so an investor who is long a 10-year bond might wish to sell short a similar 10-year bond as a hedge against it. Similarly a short position in a bond will be hedged through a purchase of an equivalent amount of the hedging instrument. In fact there are a variety of hedging instruments available, both on and off-balance sheet. Once the hedge is put on, any loss in the primary position should in theory be offset by a gain in the hedge position, and vice versa. The objective of a hedge is to ensure that the price change in the primary instrument is equal to the price change in the hedging instrument. If we are hedging a position with another bond, we use the BPVs of each bond to calculate the amount of the hedging instrument required. This is important because each bond will have different BPVs, so that to hedge a long position in say £1 million nominal of a 30-year bond does not mean simply to sell £1 million of another 30-year bond. This is because the BPVs of the two bonds will almost certainly be different. Also there might not be another 30-year bond of that particular type.

What if we have to hedge with a 10-year bond? How much nominal of this bond would be required? We need to know the ratio given at (5.19) to calculate the nominal hedge position.

$$\frac{BPV_p}{BPV_h} \tag{5.19}$$

where

BPV_p is the basis point value of the primary bond (the position to be hedged)
BPV_h is the basis point value of the hedging instrument.

The *hedge ratio* is used to calculate the size of the hedge position, and is given at (5.20).

$$\frac{BPV_p}{BPV_h} \times \frac{\text{Change in yield for primary bond position}}{\text{Change in yield for hedge instrument}} \tag{5.20}$$

The second ratio in (5.20) is known as the *yield beta*.

Example 5.2: The nature of the modified duration approximation

Table 5.1 shows the change in price for a hypothetical bond, the 8% 2009, for a selection of yields. We see that for a one basis point (bp) change in yield, the change in price given by the dollar duration figure, while not completely accurate, is a reasonable estimation of the actual change in price. For a large move however, say 200 bps, the approximation is significantly in error and analysts would not use it. Notice also for our hypothetical bond how the dollar duration value, calculated from the modified duration measurement, underestimates the change in price resulting from a fall in yields but overestimates the price change for a rise in yields. This is a reflection of the price/yield relationship for this bond. Some bonds will have a more pronounced convex relationship between price and yield, and the modified duration calculation will underestimate the price change resulting from both a fall or a rise in yields.

Convexity

Duration can be regarded as a first-order measure of interest rate risk: it measures the *slope* of the present value/yield profile. It is however only an approximation of the actual change in bond price given a small change in yield to maturity. This is also true of modified duration, which describes the price sensitivity of a bond to small changes in yield. However as Figure 5.1 illustrates, the approximation is an underestimate of the actual price at the new yield because the price/yield relationship is not linear for even plain vanilla instruments. This is the weakness of the duration measure. We see that the long-dated gilt has a reasonably convex profile, but that the callable bond and the mortgage-backed bond have slightly concave profiles, a feature known as *negative convexity*. The modified duration measure is a reasonable approximation for bonds with fixed coupon payments and maturity date, but inadequate for bonds that exhibit uncertainties in cash flow and maturity. For any bond, modified duration becomes increasingly inaccurate for increasing

Table 5.1 Nature of the modified duration approximation

Bond	*Maturity (years)*	*Modified duration*	*Price duration of basis point*	*Yield 6.00%*	*6.50%*	*7.00%*	*7.50%*	*7.99%*	***8.00%***	*8.01%*	*8.50%*	*9.00%*	*10.00%*
8% 2009	10	6.76695	0.06936	114.72017	110.78325	107.02358	103.43204	100.0671311	**100.00000**	99.932929	96.71933	93.58234	87.71087

Yield change	*Price change*	*Estimate using price duration*
Down 1 bp	0.06713	0.06936
Up 1 bp	0.06707	0.06936
Down 200 bp	14.72017	13.872
Up 200 bp	12.28913	13.872

Term (interest periods) Spot Rate

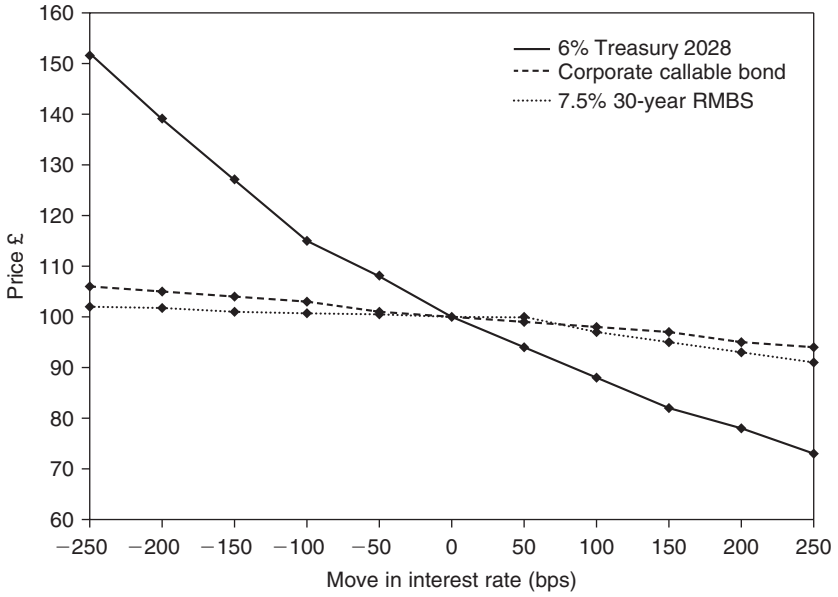


Figure 5.1 lllustration of price sensitivity for three types of bonds, 15 December 2000

magnitudes of interest rate change. For this reason, an adjustment to stimation of price change is made using the convexity measure.

Convexity is a second-order measure of interest rate risk; it measures the curvature of the present value/yield profile. Convexity can be regarded as an indication of the error we make when using duration and modified duration, as it measures the degree to which the curvature of a bond's price/yield relationship diverges from the straight-line estimation. The convexity of a bond is positively related to the dispersion of its cash flows, thus other things being equal, if one bond's cash flows are more spread out in time than another's, it will have a higher *dispersion* and hence a higher convexity. Convexity is also positively related to duration.

The second-order differential of the bond price equation with respect to the redemption yield r is:

$$\begin{aligned}\frac{\Delta P}{P} &= \frac{1}{P}\frac{\Delta p}{\Delta r}(\Delta r) + \frac{1}{2P}\frac{\Delta^2 P}{\Delta r^2}(\Delta r)^2 \\ &= -MD(\Delta r) + \frac{CV}{2}(\Delta r)^2\end{aligned} \tag{5.21}$$

where CV is the convexity.

From equation (5.21), convexity is the rate at which price variation to yield changes with respect to yield. That is, it describes a bond's modified duration

changes with respect to changes in yield. It can be approximated by expression (5.22):

$$CV = 10^8\left(\frac{\Delta P'}{P} + \frac{\Delta P''}{P}\right) \tag{5.22}$$

where

$\Delta P'$ is the change in bond price if yield increases by 1 basis point (0.01)
$\Delta P''$ is the change in bond price if yield decreases by 1 basis point (0.01).

Appendix 5.1 provides the mathematical derivation of the formula.

Example 5.3

A 5% annual coupon bond is trading at par with three years to maturity. If the yield increases from 5.00 to 5.01%, the price of the bond will fall (using the bond price equation) to:

$$P_d' = \frac{5}{(0.0501)}\left[1 - \frac{1}{(1.0501)^3}\right] + \frac{100}{(1.0501)^3} = 99.97277262$$

or by $\Delta P_d' = -0.02722738$. If the yield falls to 4.99%, the price of the bond will rise to

$$P_d' = \frac{5}{(0.0499)}\left[1 - \frac{1}{(1.0499)^3}\right] + \frac{100}{(1.0499)^3} = 100.027237$$

or by $\Delta P_d'' = 0.02723695$. Therefore

$$CV = 10^8\left(\frac{-0.02722738}{100} + \frac{0.02723695}{100}\right) = 9.57$$

That is, a convexity value of approximately 9.57.

The unit of measurement for convexity using (5.22) is the number of interest periods. For annual coupon bonds this is equal to the number of years; for bonds paying coupon on a different frequency we use (5.23) to convert the convexity measure to years.

$$CV_{years} = \frac{CV}{C^2} \tag{5.23}$$

The convexity measure for a zero-coupon bond is given by (5.24):

$$CV = \frac{n(n+1)}{(1+r)^2} \tag{5.24}$$

Convexity is a second order approximation of the change in price resulting from a change in yield. This is given by:

$$\Delta P = \tfrac{1}{2} \times CV \times (\Delta r)^2 \tag{5.25}$$

The reason we multiply the convexity by ½ to obtain the convexity adjustment is that the second term in the Taylor expansion contains the coefficient ½. The convexity approximation is obtained from a Taylor expansion of the bond price formula. An illustration of the Taylor expansion of the bond price/yield equation is given in Appendix 5.2.

The formula is the same for a semi-annual coupon bond.

Note that the value for convexity given by the expressions above will always be positive, that is, the approximate price change due to convexity is positive for both yield increases and decreases.

Example 5.4: Second-order interest rate risk

A 5% annual coupon bond is trading at par with a modified duration of 2.639 and convexity of 9.57. If we assume a significant market correction and yields rise from 5% to 7%, the price of the bond will fall by:

$$\begin{aligned}\Delta P_d &= -MD \times (\Delta r) \times P_d + \frac{CV}{2} \times (\Delta r)^2 \times P_d \\ &= -(2.639) \times (0.02) \times 100 + \frac{9.57}{2} \times (0.02)^2 \times 100 \\ &= -5.278 + 0.1914 \\ &= -£5.0866\end{aligned}$$

to £94.9134. The first-order approximation, using the modified duration value of 2.639, is −£5.278, which is an overestimation of the fall in price by £0.1914.

Example 5.5: Convexity effect

The 5% 2009 bond is trading at a price of £96.23119 (a yield of 5.50%) and has precisely 10 years to maturity. If the yield rises to 7.50%, a change of 200 basis points, the percentage price change due to the convexity effect is given by:

$$(0.5) \times 96.23119 \times (0.02)^2 \times 100 = 1.92462\%$$

If we use an HP calculator to find the price of the bond at the new yield of 7.50% we see that it is £82.83980, a change in price of 13.92%. The convexity

measure of 1.92462% is an approximation of the error we would make when using the modified duration value to estimate the price of the bond following the 200 basis point rise in yield.

If the yield of the bond were to fall by 200 basis points, the convexity effect would be the same, as given by the expression at (5.25).

Convexity is an attractive property for a bond to have. What level of premium will be attached to a bond's higher convexity? This is a function of the current yield levels in the market as well as market volatility. Remember that modified duration and convexity are functions of yield level, and that the effect of both is magnified at lower yield levels. As well as the relative level, investors will value convexity higher if the current market conditions are volatile. Remember that the cash effect of convexity is noticeable only for large moves in yield. If an investor expects market yields to move only by relatively small amounts, she will attach a lower value to convexity; and vice versa for large movements in yield. Therefore the yield premium attached to a bond with higher convexity will vary according to market expectations of the future size of interest rate changes.

The convexity measure increases with the square of maturity, and decreases with both coupon and yield. As the measure is a function of modified duration, index-linked bonds have greater convexity than conventional bonds. We discussed how the price/yield profile will be more convex for a bond of higher convexity, and that such a bond will outperform a bond of lower convexity whatever happens to market interest rates. High convexity is therefore a desirable property for bonds to have. In principle a more convex bond should fall in price less than a less convex one when yields rise, and rise in price more when yields fall. That is, convexity can be equated with the potential to outperform. Thus other things being equal, the higher the convexity of a bond the more desirable it should be in principle to investors. In some cases investors may be prepared to accept a bond with a lower yield in order to gain convexity. We noted also that convexity is in principle of more value if uncertainty, and hence expected market volatility, is high, because the convexity effect of a bond is amplified for large changes in yield. The value of convexity is therefore greater in volatile market conditions.

For a conventional vanilla bond convexity is almost always positive. Negative convexity resulting from a bond with a concave price/yield profile would not be an attractive property for a bondholder; the most common occurrence of negative convexity in the cash markets is with callable bonds.

We illustrated that for most bonds, and certainly when the convexity measure is high, the modified duration measurement for interest rate risk becomes more inaccurate for large changes in yield. In such situations it becomes necessary to use the approximation given by our convexity equation, to measure the error we have made in estimating the price change based on modified duration only. The expression was given earlier in this chapter.

The following points highlight the main convexity properties for conventional vanilla bonds.

A fall in yields leads to an increase in convexity. A decrease in bond yield leads to an increase in the bond's convexity; this is a property of positive convexity. Equally a rise in yields leads to a fall in convexity.

For a given term to maturity, higher coupon results in lower convexity. For any given redemption yield and term to maturity, the higher a bond's coupon, the lower its convexity. Therefore among bonds of the same maturity, zero-coupon bonds have the highest convexity.

For a given modified duration, higher coupon results in higher convexity. For any given redemption yield and modified duration, a higher coupon results in a higher convexity. Contrast this with the earlier property: in this case, for bonds of the same modified duration, zero-coupon bonds have the lowest convexity.

The basic redemption yield, modified duration and convexity measures are unsuitable for bond instruments that exhibit uncertain cash flow and maturity characteristics, and other techniques must be used in the analysis of such products. One such technique is the *option-adjusted spread* model, which is considered in the next chapter.

Appendix 5.1: Measuring convexity

The modified duration of a plain vanilla bond is:

$$MD = \frac{D}{(1+r)} \tag{5.26}$$

We know that:

$$\frac{dP}{dr}\frac{1}{P} = -MD \tag{5.27}$$

This shows that for a percentage change in the yield we have an inverse change in the price by the amount of the modified duration value.

If we multiply both sides of (5.27) by any particular change in the bond yield, given by *dr*, we obtain expression (5.28):

$$\frac{dP}{P} = -MD \times dr \tag{5.28}$$

Using the first two terms of a Taylor expansion, we obtain an approximation of the bond price change, given by (5.29):

$$dP = \frac{dP}{dr}dr + \frac{1}{2}\frac{d^2P}{dr^2}(dr)^2 + \text{approximation error} \tag{5.29}$$

If we divide both sides of (5.29) by P to obtain the percentage price change, the result is the expression at (5.30):

$$dP = \frac{dP}{dr}\frac{1}{P}dr + \frac{1}{2}\frac{d^2P}{dr^2}\frac{1}{P}(dr)^2 + \frac{\text{approximation error}}{P} \tag{5.30}$$

The first component of the right-hand side of (5.29) is the expression at (5.28), which is the cash price change given by the duration value. Therefore equation (5.29) is the approximation of the price change. Equation (5.30) is the approximation of the price change as given by the modified duration value. The second component in both expressions is the second derivative of the bond price equation. This second derivative captures the convexity value of the price/yield relationship, and is the cash value given by convexity. As such it is referred to as *dollar convexity* in the US markets. The dollar convexity is stated as (5.31):

$$CV_{dollar} = \frac{d^2P}{dr^2} \tag{5.31}$$

If we multiply the dollar convexity value by the square of a bond's yield change, we obtain the approximate cash value change in price resulting from the convexity effect. This is shown by (5.32):

$$dP = (CV_{dollar})(dr)^2 \tag{5.32}$$

If we then divide the second derivative of the price equation by the bond price, we obtain a measure of the percentage change in bond price as a result of the convexity effect. This is the measure known as *convexity*, and is the convention used in virtually all bond markets. This is given by the expression at (5.33):

$$CV = \frac{d^2P}{dr^2}\frac{1}{P} \tag{5.33}$$

To measure the amount of the percentage change in bond price as a result of the convex nature of the price/yield relationship we can use (5.34):

$$\frac{dP}{P} = \tfrac{1}{2}CV(dr)^2 \tag{5.34}$$

For long-hand calculations note that the second derivative of the bond price equation is (5.35), which can be simplified to (5.36). The usual assumptions apply to the expressions, that the bond pays annual coupons and has a precise number of interest periods to maturity. If the bond is a semi-annual paying one the yield value r is replaced by $r/2$.

$$\frac{d^2P}{dr^2} = \sum_{n=1}^{N}\frac{n(n+1)C}{(1+r)^{n+2}}\frac{n(n+1)M}{(1+r)^{n+2}} \tag{5.35}$$

Alternatively we differentiate to the second order the bond price equation as given by (5.36), giving us the alternative expression (5.37):

$$P = \frac{C}{r}\left[1 - \frac{1}{(1+r)^n}\right] + \frac{100}{(1+r)^n} \tag{5.36}$$

$$\frac{d^2P}{dr^2} = \frac{2C}{r^3}\left[1 - \frac{1}{(1+r)^n}\right] - \frac{2C}{r^2(1+r)^{n+1}} + \frac{n(n+1)(100 - \frac{c}{r})}{(1+r)^{n+2}} \tag{5.37}$$

Appendix 5.2: Taylor expansion of the price/yield function

Let us summarise the bond price formula as (5.38) where C represents all the cash flows from the bond, including the redemption payment:

$$P = \sum_{n=1}^{N} \frac{C_n}{(1+r)^n} \tag{5.38}$$

We therefore derive the following:

$$\frac{dP}{dr} = \sum_{n=1}^{N} \frac{C_n \bullet n}{(1+r)^{n+1}} \tag{5.39}$$

$$\frac{d^2P}{dr^2} = \sum_{n=1}^{N} \frac{C_n \bullet n(n+1)}{(1+r)^{n+2}} \tag{5.40}$$

This then gives us:

$$\Delta P = \left[\frac{dP}{dr}\Delta r\right] + \left[\frac{1}{2!}\frac{d^2P}{dr^2}(\Delta r)^2\right] + \left[\frac{1}{3!}\frac{d^3P}{dr^3}(\Delta r)^3\right] + \ldots \tag{5.41}$$

The first expression in (5.41) is the modified duration measure, while the second expression measures convexity. The more powerful the changes in yield, the more expansion is required to approximate the change to greater accuracy. Expression (5.41) therefore gives us the equations for modified duration and convexity, shown by (5.42) and (5.43) respectively.

$$MD = -\frac{dP/dr}{P} \tag{5.42}$$

$$CV = \frac{d^2P/dr^2}{P} \tag{5.43}$$

We can therefore state the following:

$$\frac{\Delta P}{P} = \left[-(MD)\Delta r\right] + \left[\tfrac{1}{2}(CV)(\Delta r)^2\right] + \text{residual error} \tag{5.44}$$

$$\Delta P = -\left[P(MD)\Delta r\right] + \left[\frac{P}{2}(CV)(\Delta r)^2\right] + \text{residual error} \tag{5.45}$$

Selected bibliography and references

Bierwag, G. O. 'Immunization, duration and the term structure of interest rates,' *Journal of Financial and Quantitative Analysis*, December 1977, pp. 725–41.

Bierwag, G. O. 'Measures of duration,' *Economic Inquiry*, 16, October 1978, pp. 497–507.

Blake, D. and Orszag, M. 'A closed-form formula for calculating bond convexity', *Journal of Fixed Income*, 6, 1996, pp. 88–91.

Burghardt, G. *The Treasury Bond Basis*, McGraw-Hill, 1994, ch. 2.

Butler, C. *Mastering Value-at-Risk*, FT Prentice Hall, 1998.

Dattatreya, R. *Fixed Income Analytics: State-of-the-Art Debt Analysis and Valuation Modeling*, McGraw-Hill, 1991.

Garbade, K. *Fixed Income Analytics*, MIT Press, 1996, chs 3, 4 and 12.

Golub, B. and Tilman, L. *Risk Management: Approaches in Fixed Income Markets*, Wiley, 2000, ch. 2.

Macaulay, F. *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields and Stock Prices in the United States since 1865*, National Bureau of Economic Research, NY, 1938 (available from Risk Classics Publishing, under the title *Interest Rates, Bond Yields and Stock Prices in the United States since 1856*).

Macaulay, F. *The Movements of Interest Rates, Bond Yields and Stock Prices in the United States Since 1856*, RISK Classics Library, 1999.

Windas, T. *An Introduction to Option-Adjusted Spread Analysis*, Bloomberg, 1993.

6

Fixed Income Securities III: Credit Analysis, Bond Spreads and Relative Value Measurement

The risks associated with holding a fixed-interest debt instrument are closely connected with the ability of the issuer to maintain the regular coupon payments as well as redeem the debt on maturity. Essentially the *credit risk* is the main risk of holding a bond. Only the highest-quality government debt, and a small number of supra-national issues, may be considered to be entirely free of credit risk. Therefore, at any time, the yield on a bond reflects investors' views on the ability of the issuer to meet its liabilities as set out in the bond's terms and conditions. A delay in paying a cash liability as it becomes due is known as technical default and is a cause for extreme concern for investors; failure to pay will result in the matter being placed in the hands of the court as investors seek to recover their funds. To judge the ability of an issue to meet its obligations for a particular debt issue, for the entire life of the issue, requires judgemental analysis of the issuer's financial strength and business prospects. There are a number of factors that must be considered; and larger banks, fund managers and corporates carry out their own *credit analysis* of individual borrowers' bond issues. The market also makes a considerable use of formal *credit ratings* that are assigned to individual bond issues by a formal credit rating agency. In the international markets the most influential ratings agencies are Standard & Poor's Corporation (S&P), Moody's Investors Service, Inc (Moody's) and Fitch Investors Service, Inc (Fitch).

The specific factors that are considered by a ratings agency, and the methodology used in conducting the analysis, differ slightly among the individual ratings agencies. Although in many cases the ratings assigned to a particular issue by different agencies are the same, they occasionally differ and in these instances investors usually seek to determine what aspect of an issuer is given more weight in each agency's analysis. Note that a credit rating is not a recommendation to buy (or equally, sell) a particular bond, nor is it a comment on market expectations. Credit analysis does take into account general market and economic conditions; the overall purpose of credit analysis is to consider the financial health of the issuer and its ability to meet the obligations of the specific issue being rated. Credit ratings play a large part in the decision-making of investors, and also have a significant impact on the interest rates payable by borrowers.

In this chapter we review credit ratings and their function, and then go on to consider the main factors involved in corporate bond credit analysis. The second part of this chapter looks at measuring bond returns and spreads.

Credit ratings

A credit rating is a formal opinion given by a rating agency, of the *credit risk* for investors in a particular issue of debt securities. Ratings are given to public issues of debt securities by any type of entity; including governments, banks and corporates. They are also given to short-term debt such as commercial paper as well as bonds and medium-term notes.

Purpose of credit ratings

Investors in securities accept the risk that the issuer will default on coupon payments or fail to repay the principal in full on the maturity date. Generally, credit risk is greater for securities with a long maturity, as there is a longer period in which the issuer may potentially default. For example, if a company issues ten-year bonds, investors cannot be certain that the company will still exist in ten years' time. It may have failed and gone into liquidation some time before that. That said, there is also risk attached to short-dated debt securities – indeed there have been instances of default by issuers of commercial paper, which is a very short-term instrument.

The prospectus or offer document for an issue provides investors with some information about the issuer so that some credit analysis can be performed on the issuer before the bonds are placed. The information in the offer documents enables investors themselves to perform their own credit analysis by studying this information before deciding whether or not to invest. Credit assessments take up time, however, and also require the specialist skills of credit analysts. Large institutional investors do in fact employ such specialists to carry out credit analysis. However, it is often too costly and time-consuming to assess every issuer in every debt market, and so investors commonly employ two other methods when making a decision on the credit risk of debt securities:

- name recognition
- formal credit ratings.

Name recognition is when the investor relies on the good name and reputation of the issuer and accepts that the issuer is of such good financial standing, or sufficient financial standing, that a default on interest and principal payments is highly unlikely. An investor may feel this way about, say, Microsoft or British Petroleum plc. However, the experience of Barings in 1995 suggested to many investors that it may not be wise to rely on name recognition alone in today's marketplace. The tradition and reputation behind the Barings name allowed the bank to borrow at Libor and occasionally at sub-Libor interest rates in the money

markets, which put it on a par with the highest-quality clearing banks in terms of credit rating. However, name recognition needs to be augmented by other methods to reduce the risk against unforeseen events, as happened with Barings. Credit ratings are a formal assessment, for a given issue of debt securities, of the likelihood that the interest and principal will be paid in full and on schedule. They are increasingly used to make investment decisions about corporate or lesser-developed government debt.

Formal credit ratings

Credit ratings are provided by the specialist agencies. On receipt of a formal request, the credit rating agencies will carry out a rating exercise on a specific issue of debt capital. The request for a rating comes from the organisation planning the issue of bonds. Although ratings are provided for the benefit of investors, the issuer must bear the cost. However, it is in the issuer's interest to request a rating as it raises the profile of the bonds, and investors may refuse to buy paper that is not accompanied with a recognised rating. Although the rating exercise involves a credit analysis of the issuer, the rating is applied to a specific debt issue. This means that, in theory, the credit rating is applied not to an organisation itself, but to specific debt securities that the organisation has issued or is planning to issue. In practice, it is common for the market to refer to the creditworthiness of organisations themselves in terms of the rating of their debt. A highly rated company such as Rabobank is therefore referred to as a 'triple-A rated' company, although it is the bank's debt issues that are rated as triple-A.

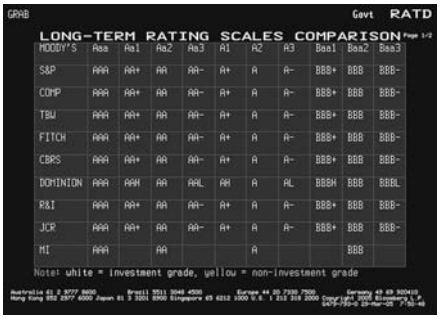
The rating for an issue is kept constantly under review and if the credit quality of the issuer declines or improves, the rating will be changed accordingly. An agency may announce in advance that it is reviewing a particular credit rating, and may go further and state that the review is a precursor to a possible downgrade or upgrade. This announcement is referred to as putting the issue under *credit watch*. The outcome of a credit watch is in most cases likely to be a rating downgrade; however, the review may reaffirm the current rating, or possibly upgrade it. During the credit watch phase the agency will advise investors to use the current rating with caution. When an agency announces that an issue is under credit watch, the price of the bonds will fall in the market as investors look to sell out of their holdings. This upward movement in yield will be more pronounced if an actual downgrade results. For example, in the beginning of 2009 the government of Spain was placed under credit watch by S&P and subsequently lost its AAA credit rating; as a result there was an immediate and sharp rise in Spain sovereign-name credit default swap premiums, before S&P had announced the actual results of its credit review (which did indeed downgrade the country risk to AA+).

A summary of long-term ratings is presented in Table 6.1.

Ratings can be accessed on the Bloomberg system. A composite page is shown at Figure 6.1, the screen RATD.

Table 6.1 Summary of credit rating agency bond ratings

Fitch	*Moody's*	*Standard & Poors*	*Summary description*
Investment grade – High credit quality			
AAA	Aaa	AAA	Gilt edged, prime, lowest-risk, risk-free
AA+	Aa1	AA+	
AA	Aa2	AA	High-grade, high credit quality
AA−	Aa3	AA−	
A+	A1	A+	
A	A2	A	Upper-medium grade
A−	A3	A−	
BBB+	Baa1	BBB+	
BBB	Baa2	BBB	Lower-medium grade
BBB−	Baa3	BBB−	
Speculative – Lower credit quality			
BB+	Ba1	BB+	
BB	Ba2	BB	Low grade; speculative
BB−	Ba3	BB−	
B+	B1	B+	
B	B2	B	Highly speculative
B−	B3	B−	
Highly speculative, substantial risk or in default			
		CCC+	
CCC	Caa	CCC	Considerable risk, in poor standing
		CCC−	
CC	Ca	CC	May already be in default, very speculative
C	C	C	Extremely speculative
		CI	Income bonds – no interest being paid
DDD			
DD			Default
D		D	



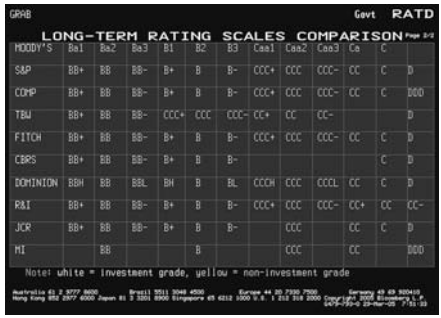


Figure 6.1 Bloomberg screen RATD.

Credit analysis

When ratings agencies were first set up, credit analysis focused primarily on the default risk of the bond, or the probability that the investor would not receive the interest payments and the principal repayment as they fell due. Although this is still important, credit analysts these days also consider the overall economic conditions, as well as the chance that an issuer will have its rating changed during the life of the bond. There are differences in approach depending on which industry or market sector the issuing company is part of.

In this section we review the main issues that are of concern to a credit analyst when rating bond issues. Analysts usually adopt a 'top-down' approach, or a 'big picture' approach, and concentrate on the macro-issues first before looking in detail at the points specific to the issuer. The process therefore involves reviewing the issuer's industry, before looking at its financial and balance sheet strength, and finally the legal provisions concerning the bond issue, as discussed below. There are also detail differences in analysis depending on which industry the issuer is in.

The issuer industry

In the first instance the credit analysis process of a specific issue will review the issuer's industry, in order to place the subsequent company analysis in context. For example, a company that has recorded growth rates of 10% each year may appear to be a quality performer, but not if its industry has been experiencing average growth rates of 30%. Generally, the industry analysis will review the following issues:

- **Economic cycle:** The business cycle of the industry and its correlation with the overall business cycle are key indicators. That is, how closely does the industry follow the rate of growth of its country's GNP? Certain industries, such as the electricity and food retail sectors, are more resistant to recession than others. Other sectors are closely tied to changes in population and birth patterns, such as residential homes, while the financial services industry is influenced by the overall health of the economy as well as by the level of interest rates. In addition to the correlation with macro-factors, credit analysts review traditional financial indicators in context, for example the issuing company's *earnings per share* (EPS) against the growth rate of its industry.
- **Growth prospects:** This review is of the issuer industry's general prospects. A company operating within what is considered a high-growth industry is generally deemed to have better credit quality expectations than one operating in a low-growth environment. A scenario of anticipated growth in the industry has implications for the issuing company, for example affecting the extent to which the company will be able cope with capacity demands and the financing of excess capacity. A fast-growth industry also attracts new entrants, which will lead to oversupply, intensified competition and reduced margins. A slow-growth industry has implications for diversification, so that a company deemed to have plans for diversifying when operating in stagnant markets will be marked up.

- **Competition:** A review of the intensity of competitive forces within an industry, and the extent of pricing and over- or under-capacity, is an essential ingredient of credit analysis. Competition is now regarded as a global phenomenon and well-rated companies are judged able to compete successfully on a global basis while concentrating on the highest-growth regions. Competition within a particular industry is related to that industry's structure and has implications for pricing flexibility. The type of market – for example, monopoly, oligopoly, and so on – also influences pricing policy and relative margins. Another issue arises if there is obvious overcapacity in an industry; this has been exemplified in the past by the airline industry and (in some countries) financial services. Overcapacity often leads to intense price competition and price wars. This is frequently damaging for the industry as a whole, as all companies suffer losses and financial deterioration in the attempt to maintain or grow market share.
- **Supply sources:** The availability of suppliers in an industry can affect a company's financial well-being. Monopoly sources of supply are considered a restrictive element, and have negative implications. A vertically integrated company that is able to supply its own raw materials is less susceptible to economic conditions that might affect suppliers or leave it hostage to price rises. A company that is not self-sufficient in its factors of production, but is nevertheless in a strong enough position to pass on its costs, is in a good position.
- **Research and development:** A broad assessment of the growth prospects of a company must also include a review of the company's research and development (R&D) position. In certain industries – such as telecommunications, media and information technology – a heavy investment in R&D is essential simply in order to maintain market share. In a high-technology field it is common for products to obsolesce very quickly, making it essential to maintain high R&D spend. In the short term, however, a company with a low level of research expenditure may actually post above-average profits (relative to the industry) because it is operating at higher margins. This is not considered a healthy strategy for the long term though.

  Evaluating the R&D input of a company is not necessarily a straightforward issue of comparing ratios, however, as it is also important to assess correctly the direction of technology. That is, a successful company needs not only to invest a sufficient amount in R&D, but must also be correct in its assessment of the direction in which the industry is heading, technology-wise. A heavy investment in developing Betamax videos, for example, would not have assisted a company in the early 1980s.
- **Level of regulation:** The degree of regulation in an industry, its direction and its effect on the profitability of a company are relevant in a credit analysis. A highly regulated industry, such as power generation, production of medicines or (in certain countries) telecommunications, can have a restrictive influence on company profits. On the other hand, if the government has announced a policy of deregulating an industry, this is considered a positive development for companies in that industry.

- **Labour relations:** An industry with a highly unionised labour force or generally tense labour relations is viewed unfavourably compared to one with stable labour relations. Credit analysts will consider historic patterns of, say, strikes and production days lost to industrial action. The status of labour relations is also more influential in a highly labour-intensive industry than in one that is more automated, for example.
- **Political climate:** The investment industry adopts an increasingly global outlook; and the emergence of sizeable tradeable debt markets in, for example, 'emerging' countries means that ratings agencies must frequently analyse the general political and economic climate in which an industry is operating. Failure to foresee certain political developments can have far-reaching effects for investors, as recently occurred in Indonesia when that country experienced a change of government; foreign investors lost funds as several local banks went bankrupt.

Financial analysis

The traditional approach to credit analysis concentrated heavily on financial analysis. The more modern approach involves first reviewing the industry within which the company is operating, as discussed above, before looking at financial considerations. Generally, the financial analysis of the issuer is conducted in three phases, namely:

- the ratio analysis for the bonds
- analysing the company's return on capital
- non-financial factors such as management expertise and extent of overseas operations.

Ratio analysis

There are a number of investor ratios that can be calculated. Ratios in themselves do not present very much insight, although there are various norms that can be applied. Generally, ratio analysis is compared to the levels prevalent in the industry, as well as historical values, in an effort to place the analysis in context and compare the company with those in its peer group. The ratios that can be considered are:

- pre-tax interest cover, which is the level of cover for interest charges in current pre-tax income
- fixed interest charge level
- *leverage*, which is commonly defined as the ratio of long-term debt as a percentage of the total capitalisation
- level of leverage compared to industry average
- nature of debt, whether fixed- or floating-rate, short- or long-term
- cash flow, which is the ratio of cash flow as a percentage of total debt. Cash flow itself is usually defined as net income from continuing operations, plus depreciation and taxes, while debt is taken to be long-term debt.
- net assets, as a percentage of total debt. The liquidity of the assets – meaning the ease with which they can be turned into cash – is taken into account when assessing the net asset ratio.

Other ratios that are considered include:

- intangibles, that is, the portion of intangibles relative to the asset side of a balance sheet
- unfunded pension liabilities. Generally, a fully funded pension is not seen as necessary, however an unfunded liability that is over 10% of net assets would be viewed as a negative point.
- age and condition of plant
- working capital.

Return on equity

There are a range of performance measures used in the market that are connected with return on equity (generally the analysis concentrates on return on capital or, more recently, risk-adjusted return on capital (RAROC)). In analysing measures of return, analysts seek to determine trends in historical performance, and make comparisons with peer group companies. Different companies also emphasise different target returns in their objectives, usually an expression of their corporate philosophy, so it is common for companies in the same industry to have different return ratios. The range of ratios used by the credit ratings agencies is shown in below. Note that 'EBIT' is 'earnings before interest and tax'.

$$\text{Return on net assets} = \frac{\text{Profit}}{\text{Net assets}} \times 100$$

$$\text{Return on sales} = \frac{\text{Profit}}{\text{Sales turnover}} \times 100$$

$$\text{Return on equity} = (\text{Return on net assets} \times \text{Gearing}) \times 100$$

$$\text{Pre-tax interest cover} = \frac{\text{Pre-tax income from continuing operations}}{\text{Gross interest}}$$

$$\text{EBIT interest cover} = \frac{\text{Pre-tax income from continuing operations} + \text{interest expense}}{\text{Gross interest}}$$

$$\text{Long-term debt as \% of capitalisation} = \frac{\text{Long-term debt}}{\text{Long-term debt} + \text{equity}} \times 100$$

$$\text{Funds flow as \% of debt} = \frac{\text{Funds from operations}}{\text{Total debt}} \times 100$$

$$\text{Free cash flow as \% of debt} = \frac{\text{Free cash flow}}{\text{Total debt}} \times 100$$

The agencies make available data that may be consulted by the public; for example, Standard & Poor's has a facility known as 'CreditStats', which was introduced in 1989. It contains the main financial ratios for a large number of companies, organised by their industry sectors.

Non-financial factors

The non-financial element of a company credit analysis has assumed a more important role in recent years, especially with regard to companies in exotic or emerging markets. Credit analysts review the non-financial factors relevant to the specific company after they have completed the financial and ratio analysis. These include the strength and competence of senior management, and the degree of exposure to overseas markets. The depth of overseas exposure is not always apparent from documents such as the annual report, and analysts sometimes need to conduct further research to determine this. Companies with considerable overseas exposure, such as petroleum companies, also need to be reviewed with respect to the political situation in their operating locations. A bank such as Standard Chartered, for example, has significant exposure to more exotic currencies in Asian, middle-eastern and African countries, and so is more at risk from additional market movements than a bank with almost exclusively domestic operations. The global, integrated nature of the bond markets also means that the foreign-exchange exposure of a company must be evaluated and assessed for risk.

The quality of management is a subjective, qualitative factor that can be reviewed in a number of ways. A personal familiarity with senior directors, acquired over a period of time, may help in the assessment. A broad breadth of experience, diversity of age, and strong internal competition between those aspiring to very senior roles, are considered positive. A company founded by one individual, in which there are no clear plans of 'succession', might be marked down.

Industry-specific analysis

Specific industries will be subject to review that is more relevant to the particular nature of the operations of the companies within them. In this section we briefly consider two separate industries, power generation and water, as well as certain other public service companies (or utilities) and financial companies.

Utility companies

The industry for power generation, water supply and (until recently) telecommunications has a tradition of being highly regulated. Until the mid-1980s, utility companies were public sector companies, and the first privatisation of such a company was British Telecom in 1984. In certain European countries utility companies are still nationalised companies, and their debt trades virtually as government debt. Credit analysis for utility companies therefore emphasises non-financial factors such as the depth of regulation and the direction in which regulation is heading, for example, towards an easing or tightening. Even in a privatised industry, new government regulation may be targeted only at the utility

sector; for example, the Labour government in the UK imposed a 'windfall tax' on several privatised utility companies shortly after being elected in May 1997.

Another consideration is government direction on how the companies may operate, such as restrictions on where a power generation company may purchase coal from – in Germany, for example, coal can be bought only from the country's own domestic coal industry, which imposes costs on the generating company that it would escape if it were free to purchase coal from other, lower-cost producers.

The financial analysis of a utility company essentially follows the pattern we described earlier.

Financial sector companies

The financial sector encompasses a large and diverse group of companies. They have an intermediary function in that they are a conduit for funds between borrowers and lenders of capital. At its simplest, a financial service company such as bank may earn profit by taking the spread between funds lent and borrowed. Banks also play an important role in managing the risk exposure for industrial companies, utilising option structures. In analysing a financial sector company the credit analyst will consider the type of customer base served by the company; for example, how much of a bank's lending is to the wholesale sector, how much is retail and so on. The financial strength and prospects of its customer base are important elements of a bank's credit rating.

Financial analysis of banks and securities houses is concerned (in addition to the factors discussed above) with the asset quality of the institution – for example, the extent of diversification of a bank's lending book. Diversification can be across customer base, as well as geographical. A loan book that is heavily concentrated in one sector is considered to be a negative factor in the overall credit assessment of the bank. A credit analyst will be concerned with the level of loans compared with levels in peer companies and the risk involved with this type of lending. For example, the expected frequency of bad loans from direct unsecured retail customer loans is higher than for retail customer loans secured by a second mortgage on a property. The higher lending rate charged for the former is designed to compensate for this higher lending risk. There are a range of financial ratios that can be used to assess a bank's asset quality. These include:

- loss reserves/net charge-off level
- net losses/average level of receivables
- non-performing loans/average level of receivables.

However, unlike the more 'concrete' financial ratios given earlier, there is a higher subjective element with these ratios as banks themselves designate which loans are non-performing and those loans against which charges have been assigned. Nevertheless, these ratios are useful indicators and may be used to identify trends across the sector as well. The loss reserves/net charge-off ratio is perhaps the most useful as it indicates the level of 'cushion' that a bank has; a falling ratio suggests that the bank may not be adding sufficient reserves to cover for

future charge-offs. This trend, if continued, may result in a future increase in the reserves and therefore a decrease in earnings levels as the expense of the reserves increases.

The leverage ratio is particularly important for financial sector companies, as the industry and business itself are highly leveraged. Banks and securities companies are therefore permitted a significantly higher leverage level than other companies. For example, in a diversified banking group with a high level of asset quality, a leverage ratio of 10:1 or even higher is considered satisfactory by ratings agencies.

Another important measure for financial companies is *liquidity*. Due to the nature of the industry and the capital structure of banks, liquidity (or more accurately the lack of liquidity) is the primary reason behind banking failures. A bank that is unable to raise funds sufficiently quickly to meet demand will most probably fail, and certainly will if external support is not provided. An inability to raise funds may arise due to internal factors, such as a deterioration in earnings, or a very poorly performing loan book, connected perhaps with a downgrade in credit rating – or from external factors, such as a major structural fault in the money markets. For credit analysis purposes the traditional liquidity measures are:

- cash
- cash equivalents
- level of receivables under one year/level of short-term liabilities.

A higher ratio indicates a greater safety cushion. A further consideration is the extent of lines of credit from other banks in the market.

Other measures of strength for financial companies are *asset coverage*, the bank's earnings record including *earnings per share* (profit attributable to shareholders/number of shares in issue) and, finally, the size of the institution. There is an element of thought which states that a very large institution, measured by asset size, cannot go bankrupt. This type of thinking can lead to complacency, however, and was one of the factors that created the financial crash of 2007–2008 (see Chapter 24).

The art of credit analysis

As bond markets become ever larger and integrated across a global market, the demand for paper is increasing – and with it the demand for high-quality credit research. There are now large numbers of companies for which investors will have virtually no recognition at all, leading to a greater reliance on formal credit ratings. Also, the rapid change in economic conditions and the effect of the business cycle frequently result in a company's credit outlook changing rapidly. Investors look primarily to credit ratings for indicators of a borrower's health. The process by which a bond issue is rated is not purely quantitative, however; and analysts frequently apply their own qualitative criteria, to take account of changing environments and other, political and macroeconomic factors.

Bond spreads and relative value

The return from a holding of fixed-income securities may be measured in more than one way. The most common approach is to consider the asset-swap spread. More sophisticated investors also consider the basis spread between the cash bond and the same-name credit default swap price, which is known the basis.[1] In this chapter we consider the most accessible way to measure bond return.

Bond spreads

Investors measure the perceived market value, or relative value, of a corporate bond by measuring its yield spread relative to a designated benchmark. This is the spread over the benchmark that gives the yield of the corporate bond. A key measure of relative value of a corporate bond is its swap spread. This is the basis point spread over the interest-rate swap curve, and is a measure of the credit risk of the bond. In its simplest form, the swap spread can be measured as the difference between the yield to maturity of the bond and the interest rate given by a straight-line interpolation of the swap curve. In practice traders use the asset-swap spread and the Z-spread as the main measures of relative value. The government bond spread is also used. In addition, now that the market in synthetic corporate credit is well established, using credit derivatives and CDS, investors consider the Cash-CDS spread as well (the basis, considered in greater detail later in this chapter).

The spread that is selected is an indication of the relative value of the bond, and a measure of its credit risk. The greater the perceived risk, the greater the spread should be. This is best illustrated by the credit structure of interest rates, which will (generally) show AAA- and AA-rated bonds trading at the lowest spreads; and BBB-, BB-rated and lower bonds trading at the highest spreads. Bond spreads are the most commonly used indication of the risk-return profile of a bond.

In this section we consider the Treasury spread, asset swap spread, Z-spread and basis.

Swap spread and Treasury spread

A bond's swap spread is a measure of the credit risk of that bond, relative to the interest-rate swaps market. Because the swaps market is traded by banks, this risk is effectively the interbank market, so the credit risk of the bond over and above bank risk is given by its spread over swaps. This is a simple calculation to make, and is simply the yield of the bond minus the swap rate for the appropriate maturity swap. Figure 6.2 shows Bloomberg page IRSB for Pounds sterling as at 10 August 2005. This shows the GBP swap curve on the left-hand side. The right-hand side of the screen shows the swap rates' spread over UK gilts. It is the spread over these swap rates that would provide the simplest relative value measure for corporate bonds denominated in GBP. If the bond has an odd maturity, say 5.5 years, we would interpolate between the 5-year and 6-year swap rates.

[1]Credit default swaps are reviewed in Chapter 20. For more details on the cash-CDS basis see Choudhry (2010).

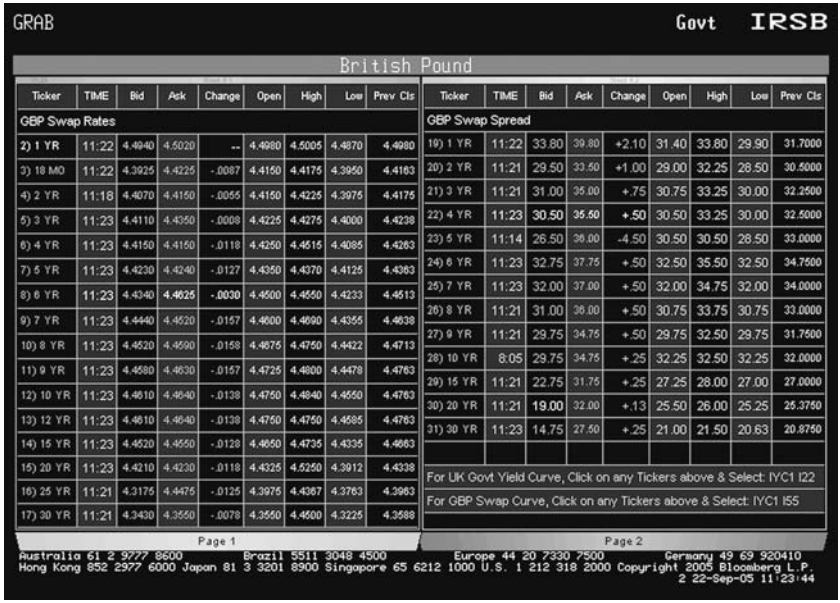


Figure 6.2 Bloomberg page IRSB for Pounds sterling, showing GBP swap rates and swap spread over UK gilts © Bloomberg L.P. All rights reserved. Reprinted with permission.

The spread over swaps is sometimes called the *I-spread*. It has a simple relationship to swaps and Treasury yields, shown here in the equation for corporate bond yield,

$$Y = I + S + T \tag{6.1}$$

where

Y is the yield on the corporate bond
I is the I-spread or spread over swap
S is the swap spread
T is the yield on the Treasury security (or an interpolated yield).

In other words, the swap rate itself is given by $T + S$.

The I-spread is sometimes used to compare a cash bond with its equivalent CDS price, but for straightforward relative value analysis it is usually dropped in favour of the asset-swap spread, which we look at later in this section.

Of course the basic relative value measure is the Treasury spread or government bond spread. This is simply the spread of the bond yield over the yield of the appropriate government bond. Again, an interpolated yield may need to be used to obtain the right Treasury rate to use. The bond spread is given by:

$$BS = Y - T.$$

Using an interpolated yield is not strictly accurate because yield curves are smooth in shape and so straight-line interpolation will produce slight errors. The method is still commonly used though.

Asset-swap spread

An asset swap is a package that combines an interest-rate swap with a cash bond, the effect of the combined package being to transform the interest-rate basis of the bond. Typically, a fixed-rate bond will be combined with an interest-rate swap in which the bond holder pays fixed coupon and receives floating coupon. The floating coupon will be a spread over Libor. This spread is the asset-swap spread and is a function of the credit risk of the bond over and above interbank credit risk.[2] Asset swaps may be transacted at par or at the bond's market price, usually par. This means that the asset swap value is made up of the difference between the bond's market price and par, as well as the difference between the bond coupon and the swap fixed rate.

The zero-coupon curve is used in the asset swap valuation. This curve is derived from the swap curve, so it is the implied zero-coupon curve. The asset swap spread is the spread that equates the difference between the present value of the bond's cash flows, calculated using the swap zero rates, and the market price of the bond. This spread is a function of the bond's market price and yield, its cash flows and the implied zero-coupon interest rates.[3]

Figure 6.3 shows the Bloomberg screen ASW for a GBP-denominated bond, GKN Holdings 7% 2012, as at 10 August 2005. We see that the asset-swap spread is 121.5 basis points. This is the spread over Libor that will be received if the bond is purchased in an asset-swap package. In essence the asset swap spread measures a difference between the market price of the bond and the value of the bond when cash flows have been valued using zero-coupon rates. The asset-swap spread can therefore be regarded as the coupon of an annuity in the swap market that equals this difference.

The Z-spread

The conventional approach for analysing an asset swap uses the bond's yield to maturity (YTM) in calculating the spread. The assumptions implicit in the YTM calculation (see Chapter 2) make this spread problematic for relative analysis, so market practitioners use what is termed the Z-spread instead. The Z-spread uses the zero-coupon yield curve to calculate spread, and so is a more realistic, and effective, spread to use. The zero-coupon curve used in the calculation is derived from the interest-rate swap curve.

---

[2]This is because in the interbank market, two banks transacting an interest-rate swap will be paying/receiving the fixed rate and receiving/paying Libor-flat. See also the 'Learning Curve' article on asset swaps available on www.yieldcurve.com.

[3]Bloomberg refers to this spread as the Gross Spread.

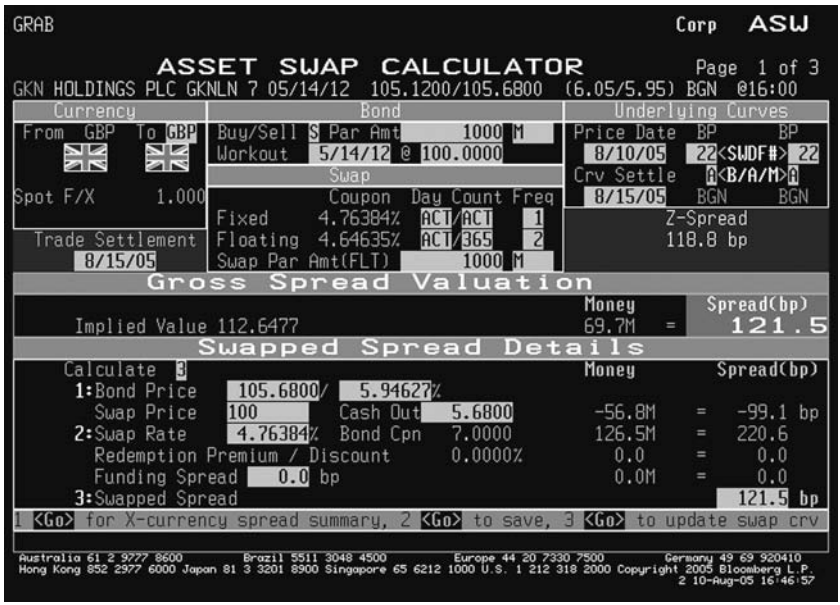


Figure 6.3 Bloomberg page ASW for GKN bond, 10 August 2005 © Bloomberg L.P. All rights reserved. Reprinted with permission.

Put simply, the Z-spread is the basis point spread that would need to be added to the implied spot yield curve such that the discounted cash flows of the bond are equal to its present value (its current market price). Each bond cash flow is discounted by the relevant spot rate for its maturity term. How does this differ from the conventional asset-swap spread? Essentially, in its use of zero-coupon rates when assigning a value to a bond. Each cash flow is discounted using its own particular zero-coupon rate. The price of a bond at any time can be taken to be the market's value of the bond's cash flows. Using the Z-spread we can quantify what the swap market thinks of this value; that is, by how much the conventional spread differs from the Z-spread. Both spreads can be viewed as the coupon of a swap market annuity of equivalent credit risk of the bond being valued.

In practice, the Z-spread – especially for shorter-dated bonds and for better credit-quality bonds – does not differ greatly from the conventional asset-swap spread. The Z-spread is usually the higher spread of the two, following the logic of spot rates, but not always. If it differs greatly, then the bond can be considered to be mis-priced.

Figure 6.4 is the Bloomberg screen YAS for the same bond shown in Figure 6.3, as at the same date. It shows a number of spreads for the bond. The main spread of 151.00 bps is the spread over the government yield curve. This is an interpolated spread, as can be seen lower down the screen, with the appropriate benchmark bind identified. We see that the asset-swap spread is 121.6 bps, while the

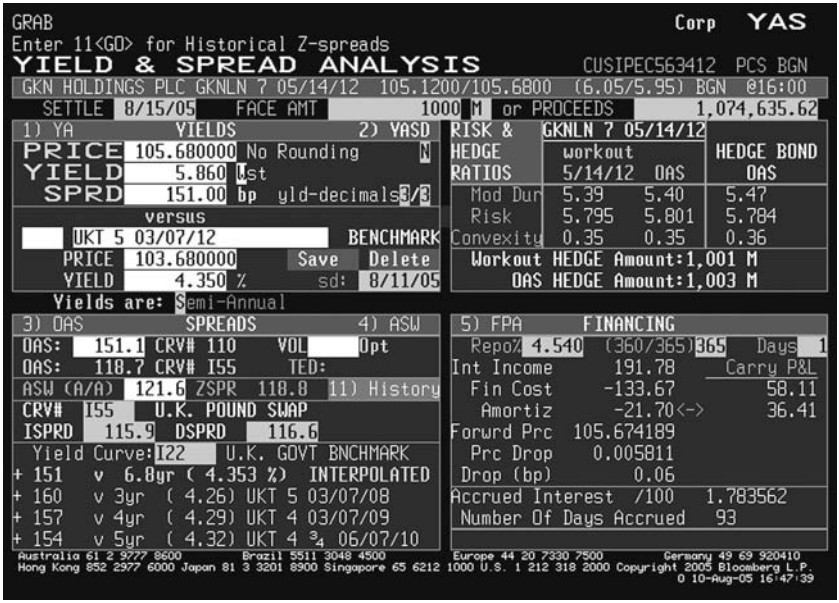


Figure 6.4 Bloomberg page YAS for GKN bond, 10 August 2005 © Bloomberg L.P. All rights reserved. Reprinted with permission.

Z-spread is 118.8 bps. When undertaking relative value analysis, for instance if making comparisons against cash funding rates or the same company name credit default swap (CDS), it is this lower spread that should be used.[4]

The same screen can be used to check spread history. This is shown at Figure 6.5, the Z-spread graph for the GKN bond for the six months prior to our calculation date.

The Z-spread is closely related to the bond price, as shown by:

$$P = \sum_{i=1}^{n} \left[\frac{C_i + M_i}{(1 + ((Z + S_i + T_i)/m))^i} \right] \tag{6.2}$$

where

n is the number of interest periods until maturity
P is the bond price
C is the coupon
M is the redemption payment (so bond cash flow is all C plus M)
Z is the Z-spread
m is the frequency of coupon payments.

[4]On the date in question the 10-year CDS for this reference entity was quoted as 96.8 bps, which is an example of a negative basis, in this case of −22 bps.

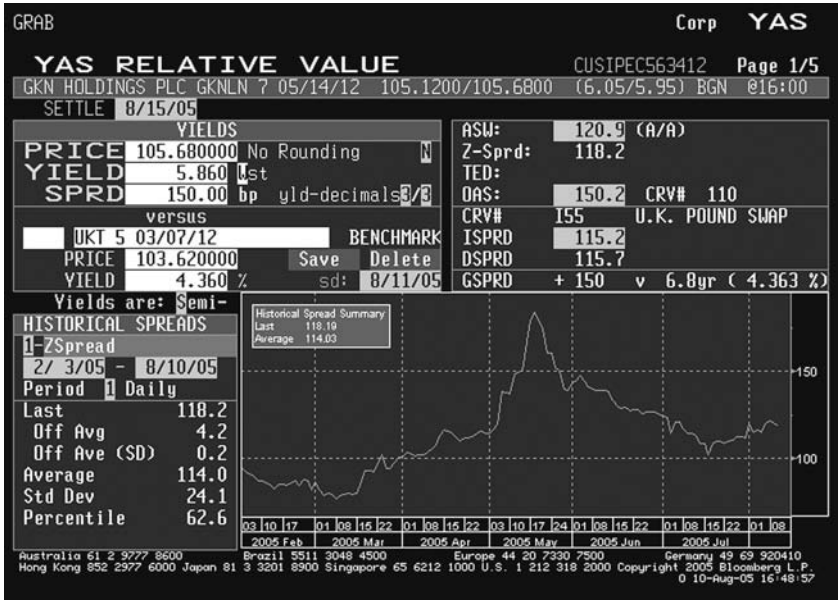


Figure 6.5 Bloomberg page YAS for GKN bond, 10 August 2005 showing Z-spread history
© Bloomberg L.P. All rights reserved. Reprinted with permission.

In effect this is the standard bond price equation with the discount rate adjusted by whatever the Z-spread is; it is an iterative calculation. The appropriate maturity swap rate is used, which is the essential difference between the I-spread and the Z-spread. This is deemed to be more accurate, because the entire swap curve is taken into account rather than just one point on it. In practice though, as we have seen in the example above, there is often little difference between the two spreads.

To reiterate then, using the correct Z-spread, the sum of the bond's discounted cash flows will be equal to the current price of the bond.

We illustrate the Z-spread calculation at Figure 6.6. This is done using a hypothetical bond, the XYZ plc 5% of June 2008, a three-year bond at the time of the calculation. Market rates for swaps, Treasury and CDS are also shown. We require the spread over the swaps curve that equates the present values of the cash flows to the current market price. The cash flows are discounted using the appropriate swap rate for each cash flow maturity. With a bond yield of 5.635%, we see that the I-spread is 43.5 basis points, while the Z-spread is 19.4 basis points. In practice the difference between these two spreads is rarely this large.

For the reader's benefit we also show the Excel formula in Figure 6.6. This shows how the Z-spread is calculated; for ease of illustration we have assumed that the calculation takes place for value on a coupon date, so that we have precisely an even period to maturity.

A1	B	C	D	E	F	G	H	I
2	**Issuer**	**XYZ plc**						
3	Settlement date	6/1/2005						
4	Maturity date	6/1/2008						
5	Coupon	5%		**YIELD**	**0.05635**			
6	Price	98.95		[Cell formula =YIELD(C4,C5,C6,C7,C8,C9,C10)]				
7	Par	100						
8	Semi-annual coupon	2		**PRICE**	**98.95000**			
9	act/act	1		[Cell formula =PRICE(C4,C5,C6,C6,C8,C9,C10)]				
10								
11	Bond yield	5.635%						
12	Sovereign bond yield	4.880%						
13	Swap rate	5.200%						
14								
15	3-year CDS price	28 bps						
16								
17	**Treasury spread**							
18	5.635 - 4.88	55 bps						
19								
20	**I-spread**							
21	5.635 - 5.20	43.5 bps						
22								
23	**Z-spread (Z)**	19.4 bps	0.00194					
24	The Z-spread is found using iteration							
25								

							Sum of PVs
Cash flow date	12/1/2005	6/1/2006	12/1/2006	6/1/2007	12/1/2007	6/1/2008	
Cash flow maturity (years)	0.50	1.00	1.50	2.00	2.50	3.00	
0.5-year swap rate (S)	4.31%	4.84%	4.99%	5.09%	5.18%	5.20%	
Cash flow (CF)	2.50	2.50	2.50	2.50	2.50	102.50	
Discount factor	0.97797598	0.951498751	0.926103469	0.900947692	0.875835752	0.852419659	
(DF Calculation)	1/(1+(S+Z)/2)^1	1/(1+(S+Z)/2)^2	1/(1+(S+Z)/2)^3	1/(1+(S+Z)/2)^4	1/(1+(S+Z)/2)^5	1/(1+(S+Z)/2)^6	
CF present value (PV)	2.445	2.379	2.315	2.252	2.190	87.373	**98.95**

34			
35			
36			
37	A Z-spread of 19.4 basis points gives us the current bond price so is the correct one		
38	Using this value, the sum of all the discounted cashflows is equal to the market price		
39			
40	**CDS Basis**		
41	28 - 19.4	8.6 bps	
42	The basis is positive in this example		

Figure 6.6 Calculating the Z-spread, hypothetical 5% 2008 bond issued by XYZ plc

Cash-CDS basis

The difference between the premium on a CDS contract and the same name (and same tenor) cash bond yield is known as the basis. That is, the basis is the CDS spread minus the ASW spread. Alternatively, it can be the CDS spread minus the Z-spread. So the basis is given by

$$B = D - CashSpread \tag{6.3}$$

where D is the CDS price. Where $D - CashSpread > 0$ it is a positive basis; the opposite is a negative basis.

Figure 6.7 shows page G <go> on Bloomberg, set up to show the Z-spread and CDS price history for the GKN 2012 bond, for the period March–September 2005. We can select the 'Table' option to obtain the actual values, which can then be used to plot the basis. This is shown at Figure 6.8, for the period 22 August to 22 September 2005. Notice how the basis was always negative during August–September; we see from Figure 6.7 that earlier in the year the basis had briefly been positive. Changes in the basis give rise to arbitrage opportunities between the cash and synthetic markets. This is discussed in greater detail in Choudhry (2006).

A wide range of factors drive the basis, which are described in detail in Choudhry (*ibid*). The existence of a non-zero basis has implications for investment strategy. For instance, when the basis is negative, investors may prefer to hold the cash

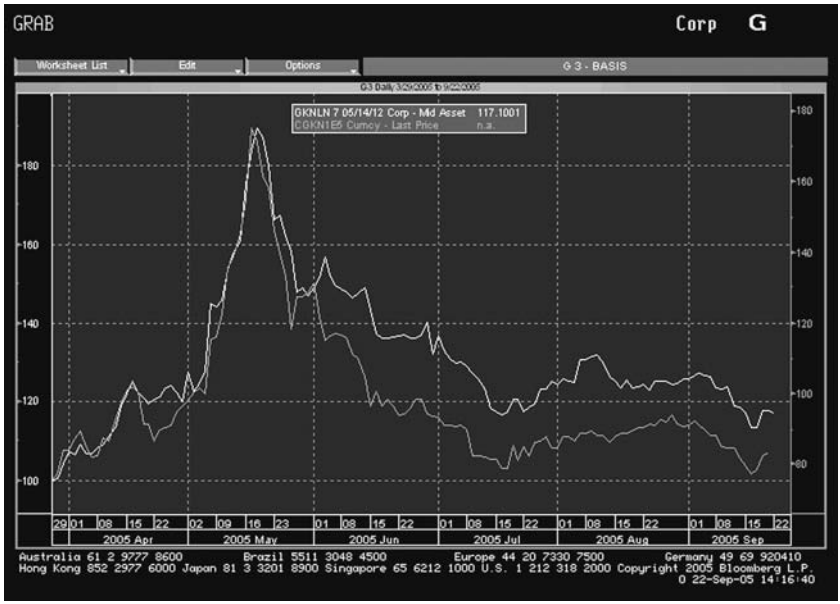


Figure 6.7 Bloomberg graph using screen G <go>, plot of asset-swap spread and CDS price for GKN bond, April–September 2005 © Bloomberg L.P. All rights reserved. Reprinted with permission.

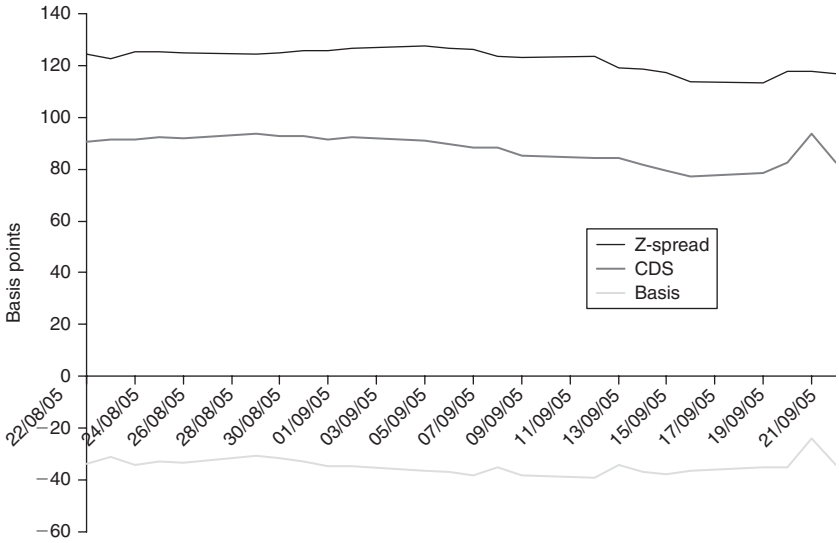


Figure 6.8 GKN bond, CDS basis during August–September 2005.
Source: Bloomberg L.P.

bond; but if (for liquidity, supply or other reasons) the basis is positive, the investor may wish to hold the asset synthetically, by selling protection using a credit default swap. Another approach is to arbitrage between the cash and synthetic markets, in the case of a negative basis; buying the cash bond and shorting it synthetically by buying protection in the CDS market. Investors have a range of spreads to use when performing their relative value analysis.

References

Choudhry, M., *The Credit Default Swap Basis*, Bloomberg Press 2006.
Choudhry, M., *Structured Credit Products: Credit Derivatives and Synthetic Securitisation*, 2nd edition, Singapore: John Wiley & Sons 2010.

7
Interest Rate Modelling

Chapter 4 introduced the concept of the yield curve. The analysis and valuation of debt market instruments revolves around the yield curve. Yield curve or *term structure* modelling has been extensively researched in the financial economics literature; it is possibly the most heavily covered subject in that field. It is not possible to deliver a comprehensive summary in just one chapter, but our aim is to cover the basic concepts. As ever, interested readers are directed to the bibliography, which lists the more accessible titles in this area.

In this chapter we review a number of interest-rate models, generally the more well-known ones. In the next two chapters we discuss some of the techniques used to fit a smooth yield curve to market-observed bond yields, and present an advanced treatment of the B-spline curve fitting methodology.

Introduction

Term structure modelling is based on theory describing the behaviour of interest rates. A model would seek to identify the elements or *factors* that are believed to explain the dynamics of interest rates. These factors are random or *stochastic* in nature, so that we cannot predict with certainty the future level of any factor. An interest-rate model must therefore specify a statistical process that describes the stochastic property of these factors, in order to arrive at a reasonably accurate representation of the behaviour of interest rates.

The first term structure models described in the academic literature described the interest-rate process as one where the *short rate*[1] follows a statistical process and where all other interest rates are a function of the short rate, so the dynamics of the short rate drive all other term interest rates. These models are known as *one-factor models*. A one-factor model assumes that all term rates follow once the short rate is specified, that is, they are not randomly determined. Two-factor interest-rate models have also been described. For instance the model described by Brennan and Schwartz (1979) specified the factors as the short rate and a long-term

[1]The short rate is a theoretical construct that refers to the interest rate that would be charged on a loan of funds that is repaid almost instantaneously.

rate, while a model described by Fong and Vasicek (1991) specified the factors as the short rate and short-rate volatility.

Basic concepts

The original class of interest-rate models described the dynamics of the short rate; the later class of models known as 'HJM' models described the dynamics of the forward rate, and we will introduce these later. The foundation of interest-rate models is grounded in probability theory, so readers may wish to familiarise themselves with this subject. An excellent introduction to this area is given in Ross (1999), while a fuller treatment is given in the same author's better-known book, *Probability Models* (2000).

In a one-factor model of interest rates, the short rate is assumed to be a random or stochastic variable, with the dynamics of its behaviour being uncertain and acting in an unpredictable manner. A random variable such as the short rate is defined as a variable whose future outcome can assume more than one possible value. Random variables are either *discrete* or *continuous*. A discrete variable moves in identifiable breaks or jumps so for example while time is continuous, the trading hours of an exchange-traded future are not continuous, as the exchange will be shut outside business hours. Interest rates are treated in academic literature as being continuous, whereas in fact rates such as central bank base rates move in discrete steps. A continuous variable moves in a manner that has no breaks or jumps, so if an interest rate can move in a range from 5% to 10%, if it is continuous it can assume any value between this range, for instance a value of 5.671291%. Although this does not reflect market reality, assuming that interest rates and the processes they follow are continuous allows us to use calculus to derive useful results in our analysis.

The short rate is said to follow a stochastic process, so although the rate itself cannot be predicted with certainty, as it can assume a range of possible values in the future, the process by which it changes from value to value can be assumed, and hence modelled. The dynamics of the short rate therefore are a stochastic process or probability distribution. A one-factor model of the interest rate actually specifies the stochastic process that describes the movement of the short rate.

The analysis of stochastic processes employs mathematical techniques originally used in physics. An instantaneous change in value of a random variable x is written as dx. The changes in the random variable are assumed to be normally distributed. The shock to this random variable that generates its change in value, also referred to as *noise*, follows a randomly generated process known as a Weiner process or geometric Brownian motion. This is described in Appendix 7.1. A variable following a Weiner process is a random variable, termed x or z, whose value alters instantaneously, but whose patterns of change follow a normal distribution with mean 0 and standard deviation 1. If we assume that the yield r of a zero-coupon bond follows a continuous Weiner process with mean 0 and standard deviation 1, this would be written:

$$dr = dz$$

Changes or 'jumps' in the yield that follow a Weiner process are scaled by the volatility of the stochastic process that drives interest rates, which is given by σ. So the stochastic process for the change in yields is given by:

$$dr = \sigma dz$$

The value of this volatility parameter is user-specified, that is, it is set at a value that the user feels most accurately describes the current interest-rate environment. Users often use the volatility implied by the market price of interest-rate derivatives such as caps and floors.

So far we have said that the zero-coupon bond yield is a stochastic process following a geometric Brownian motion that drifts with no discernible trend; however under this scenario, over time the yield would continuously rise to a level of infinity or fall to infinity, which is not an accurate representation of reality. We need to add to the model a term that describes the observed trend of interest rates moving up and down in a cycle. This expected direction of the change in the short rate is the second parameter in an interest-rate model, which in some texts is referred to by a letter such as a or b and in other texts is referred to as μ.

The short-rate process can therefore be described in the functional form given by (7.1):

$$dr = adt + \mu dz \tag{7.1}$$

where

dr is the change in the short rate
a is the expected direction of change of the short rate or drift
dt is the incremental change in time
μ is the standard deviation of changes in the short rate
dz is the random process.

Equation (7.1) is sometimes seen with dW or dx in place of dz. It assumes that on average the instantaneous change in interest rates is given by the function adt, with random shocks specified by σdz. It is similar to a number of models, such as those first described by Vasicek (1977), Ho and Lee (1986), Hull and White (1990) and others.

To reiterate, (7.1) states that the change in the short rate r over an infinitesimal period of time dt, termed dr, is a function of:

- the drift rate or expected direction of change in the short rate a
- a random process dz.

The two significant properties of the geometric Brownian motion are:

- The drift rate is equal to the expected value of the change in the short rate. Under a zero drift rate, the expected value of the change is also zero and the expected value of the short rate is given by its current value.
- The variance of the change in the short rate over a period of time T is equal to T, while its standard deviation is given by $\sqrt{T}$.

The model given by (7.1) describes a stochastic short rate process, modified with a drift rate to influence the direction of change. However a more realistic specification would also build in a term that describes the long-run tendency of interest rates to drift back to a long-run level. This process is known as *mean reversion*, and is perhaps is best known by the Hull–White model. A general specification of mean reversion would be a modification given by (4.2):

$$dr = a(b - r)\,dt + \sigma dz \tag{7.2}$$

where b is the long-run mean level of interest rates and where a now describes the speed of mean reversion. Equation (7.2) is known as an Ornstein–Uhlenbeck process. When r is greater than b, it will be pulled back towards b, although random shocks generated by dz will delay this process. When r is below b the short rate will be pulled up towards b.

Ito's lemma

Once a term structure model has been specified, it becomes necessary for market practitioners to determine how security prices related to interest rates fluctuate. The main instance of this is where we wish to determine how the price P of a bond moves over time and as the short rate r varies. The formula used for this is known as Ito's lemma. For the background on the application of Ito's lemma see Hull (1997), or Baxter and Rennie (1996). Ito's lemma transforms the dynamics of the bond price P in terms into a stochastic process in the following form:

$$dP = P_r dr + \frac{1}{2} P_{rr} (dr)^2 + P_t \tag{7.3}$$

The subscripts indicate partial derivatives.[2] The terms dr and $(dr)^2$ are dependent on the stochastic process that is selected for the short rate r. If this process is the Ornstein–Uhlenbeck process that was described in (7.2), then the dynamics of P can be specified as (7.4).

$$\begin{aligned} dP &= P_r[a(b-r)dt + \sigma dz] + \frac{1}{2} P_{rr}\sigma^2 dt + P_t dt \\ &= \left[P_r a(b-r) + \frac{1}{2} P_{rr}\sigma^2 + P_t \right] dt + P_r \sigma dz \\ &= a(r,t)dt + \sigma(r,t)dz \end{aligned} \tag{7.4}$$

What we have done is to transform the dynamics of the bond price in terms of the drift and volatility of the short rate. Equation (7.4) states that the bond price depends on the drift of the short rate, and the volatility.

[2] This is the great value of Ito's lemma, a mechanism by which we can transform a partial differential equation.

Ito's lemma is used as part of the process of building a term structure model. The generic process this follows involves the following:

- Specify the random or stochastic process followed by the short rate, for which we must make certain assumptions about the short rate itself.
- Use Ito's lemma to transform the dynamics of the zero-coupon bond price in terms of the short rate.
- Impose no-arbitrage conditions, based on the principle of hedging a position in one bond with one in another of bond of a different maturity (for a one-factor model), in order to derive the partial differential equation of the zero-coupon bond price. (For a two-factor model we would require two bonds as hedging instrument.)
- Solve the partial differential equation for the bond price, which is subject to the condition that the price of a zero-coupon bond on maturity is 1.

In the next section we review some of the models that are used in this process.

One-factor term structure models

In this section we discuss briefly a number of popular term structure models and attempt to summarise the advantages and disadvantages of each, which renders them useful or otherwise under certain conditions and user requirements.

The Vasicek model

The Vasicek model (1977) was the first term structure model described in the academic literature, and is a yield-based one-factor equilibrium model. It assumes that the short rate process follows a normal distribution. The model incorporates mean reversion and is popular with certain practitioners as well as academics because it is analytically tractable.[3] Although it has a constant volatility element, the mean reversion feature means that the model removes the certainty of a negative interest rate over the long term. However other practitioners do not favour the model because it is not necessarily arbitrage-free with respect to the prices of actual bonds in the market.

The instantaneous short rate as described in the Vasicek model is:

$$dr = a(b - r)dt + \sigma dz \tag{7.5}$$

where a is the speed of the mean reversion and b is the mean reversion level or the long-run value of r, and z is the standard Weiner process or Brownian motion with a 0 mean and 1 standard deviation. In Vasicek's model the price at time t of a zero-coupon bond that matures at time T is given by

$$P(t,T) = \mathrm{A}(t,T)e^{-B(t,T)r(t)} \tag{7.6}$$

[3] *Tractability* is much prized in a yield curve model, and refers to the ease with which a model can be implemented, that is, with which yield curves can be computed.

where $r(t)$ is the short rate at time t and

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

and

$$A(t,T) = \exp\left[\frac{(B(t,T) - T + t)(a^2 b - \sigma^2 / 2)}{a^2} - \frac{\sigma^2 B(t,T)^2}{4a}\right]$$

The derivation of (7.6) is given in a number of texts (not least the original article!). We recommend section 5.3 in Van Deventer and Imai (1997) for its accessibility.

Note that in certain texts the model is written as:

$$dr = \kappa(\theta - r)dt + \sigma dz$$

or

$$dr = \alpha(\mu - r)dt + \sigma dZ$$

but it just depends on which symbol the particular text is using. We use the form shown at (7.5) because it is consistent with discussion elsewhere in this book.

In Vasicek's model the short rate r is normally distributed, so therefore it can be negative with positive probability. The occurrence of negative rates is dependent on the initial interest rate level and the parameters chosen for the model, and is an extreme possibility. For instance a very low initial rate, such as that observed in the Japanese economy for some time now, and volatility levels set with the market, have led to negative rates when using the Vasicek model. This possibility, which also applies to a number of other interest rate models, is inconsistent with a no-arbitrage market because investors will hold cash rather than opt to invest at a negative interest rate.[4] However for most applications the model is robust, and its tractability makes it popular with practitioners.

The Ho and Lee model

The Ho and Lee model (1986) was an early arbitrage-free yield-based model. It is often called the extended Merton model as it is an extension of an earlier model described by Merton in 1970 (see Merton 1993).[5] It is called an arbitrage model as is it used to fit a given initial yield curve. The model assumes a normally distributed short rate, and the drift of the short rate is dependent on time, which makes the model arbitrage-free with respect to observed prices in the market, as these are the inputs to the model.

The model is given at (7.7).

$$dr = a(t)dt + \sigma dz \tag{7.7}$$

[4] This is stated in Black (1995).

[5] The reference in the bibliography is a later publication that is a collection of Merton's earlier papers.

The bond price equation is given as

$$P(t,T) = A(t,T)\mathrm{e}^{-\mathrm{r(t)}(T-t)} \tag{7.8}$$

where $r(t)$ is the rate at time t and

$$\ln A(t,T) = \ln\left(\frac{(P(0,T)}{(P(0,t)}\right) - (T-t)\frac{\partial \ln P(0,t)}{\partial t} - \frac{1}{2}\sigma^2(T-t)^2$$

There is no mean reversion feature incorporated so that interest rates can fall to negative levels, which is a cause for concern for market practitioners.

The Hull and White model

The model described by Hull and White (1990) is another well-known model that fits the theoretical yield curve that one would obtain using Vasicek's model extracted from the actual observed market yield curve. As such it is sometimes referred to as the extended Vasicek model, with time-dependent drift.[6] The model is popular with practitioners precisely because it enables them to calculate a theoretical yield curve that is identical to yields observed in the market, which can then be used to price bonds, bond derivatives and also calculate hedges.

The model is given at (7.9).

$$dr = a\left(\frac{b(t)-r}{a}\right)dt + \sigma dz \tag{7.9}$$

where a is the rate of mean reversion and $\frac{b(t)}{a}$ is a time-dependent mean reversion.

The price at time t of a zero-coupon bond with maturity T is

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

where $r(t)$ is the short rate at time t and

$$\ln A(t,T) = \ln\left(\frac{(P(0,T)}{(P(0,t)}\right) - B(t,T)\frac{\partial P(0,t)}{\partial t} - \frac{v(t,T)^2}{2}$$

and

$$v(t,T)^2 = \frac{1}{2a^3}\sigma^2(e^{-aT}-e^{-at})^2(e^{2at}-1)$$

Further one-factor term structure models

The academic literature and market application have thrown up a large number of term structure models as alternatives to the Vasicek model and models based on it such as the Hull–White model. As with these two models, each possesses a number

[6] Haug (1998) also states that the Hull–White model is essentially the Ho and Lee model with mean reversion.

of advantages and disadvantages. As we noted in the previous section, the main advantage of Vasicek-type models is their analytic tractability, with the assumption of the dynamics of the interest rate allowing the analytical solution of bonds and bond instruments. The main weakness of these models is that they permit the possibility of negative interest rates. While negative interest rates are not a market impossibility,[7] the thinking would appear to be that they are a function of more than one factor, therefore modelling them using Vasicek-type models is not tenable. This aspect of the models does not necessarily preclude their use in practice, which will depend on the state of the economy at the time.

To consider an example, during 1997–8 Japanese money market interest rates were frequently below 0.5%, and at this level even low levels of volatility below 5% will imply negative interest rates with high probability if Vasicek's model is used. In this environment practitioners may wish to use models that do not admit the possibility of negative interest rates, perhaps those that model more than the short rate alone, so-called two-factor and multi-factor models. We look briefly at these in the next section. First we consider, again briefly, a number of other one-factor models. As usual, readers are encouraged to review the bibliography articles for the necessary background and further detail on application.

The Cox, Ingersoll and Ross model

Although published officially in 1985, the Cox–Ingersoll–Ross model was apparently described in academic circles in 1977 or perhaps earlier, which would make it the first interest-rate model. Like the Vasicek model it is a one-factor model that defines interest rate movements in terms of the dynamics of the short rate. However its incorporates an additional feature whereby the variance of the short rate is related to the level of interest rates, and this feature has the effect of not allowing negative interest rates. It also reflects a higher interest rate volatility in periods of relatively high interest rates, and corresponding lower volatility when interest rates are lower.

The model is given at (7.10).

$$dr = k(b-r)dt + \sigma\sqrt{r}dz \tag{7.10}$$

The derivation of the zero-coupon bond price equation given at (7.11) is contained in Ingersoll (1987), chapter 18. The symbol represents the term to maturity of the bond or $(T - t)$.

$$P(r,\tau) = A(\tau)e^{-B(\tau)r} \tag{7.11}$$

[7] Negative interest rates manifest themselves most obviously in the market for specific bonds in repo which have gone excessively special. However academic researchers often prefer to work with interest rate environments that do not consider negative rates a possibility (for example see Black 1995).

where

$$A(\tau) = \left[\frac{2ye(y+\lambda+k)\frac{\tau}{2}}{g(\tau)} \right]^{\frac{2kb}{\sigma^2}}$$

$$B(\tau) = \frac{-2(1-e^{-yt})}{g(\tau)}$$

$$g(\tau) = 2\gamma + (\kappa+\lambda+y)(e^{\gamma\tau}-1)$$

$$\gamma = \sqrt{(k+\lambda)^2+\sigma^2}$$

Some researchers[8] have stated that the difficulties in determining parameters for the CIR model have limited its use among market practitioners.

The Black, Derman and Toy model

The Black–Derman–Toy model (1990) also removes the possibility of negative interest rates and is commonly encountered in the markets. The parameters specified in the model are time-dependent, and the dynamics of the short rate process incorporate changes in the level of the rate. The model is given at (7.12).

$$d[\ln(r)] = [\vartheta(t) - \phi(t)\ln(r)]dt + \sigma(t)dz \tag{7.12}$$

The popularity of the model among market practitioners reflects the following:

- It fits the market-observed yield curve, similar to the Hull–White model.
- It makes no allowance for negative interest rates.
- It models the volatility levels of interest rates in the market.

Against this, the model is not considered particularly tractable or able to be programmed for rapid calculation. Nevertheless it is important in the market, particularly for interest-rate derivative market makers. An excellent and accessible description of the BDT model is contained in Sundaresan (1997) on pages 240–4; Tuckman (1996) pages 102–6 is also recommended.

The Heath, Jarrow and Morton model

We have devoted a separate section to the approach described by Heath, Jarrow and Morton (1992) because it is a radical departure from the earlier family of interest rate models. As usual a fuller exposition can be found in the references listed in the bibliography.

[8] For instance see Van Deventer and Imai (1997) citing Fleseker (1993) on page 336, although the authors go on to state that the CIR model is deserving of further empirical analysis and remains worthwhile for practical application.

The Heath–Jarrow–Morton (HJM) approach to the specification of stochastic state variables is different from that used in earlier models. The previous models describe interest-rate dynamics in terms of the short rate as the single or (in two and multi-factor models) key state variable. With multi-factor models, the specification of the state variables is the fundamental issue in practical application of the models themselves. In the HJM model, the entire term structure and not just the short rate is taken to be the state variable. It has been seen previously how the term structure can be defined in terms of default-free zero-coupon bond prices, yields, spot rates or forward rates. The HJM approach uses forward rates. So in the single-factor HJM model the change in forward rates at current time t, with a maturity at time u, is captured by:

- a volatility function
- a drift function
- a geometric Brownian or Weiner process which describes the shocks or noise experienced by the term structure.

We present here a brief introduction to the HJM model. The account of the single-factor HJM model follows (with permission) the approach contained in chapter 5 of Baxter and Rennie (1996). This is an accessible and excellent text and is strongly recommended. Another recommended reading is James and Webber (2000).

The importance of the HJM presentation is that 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, principally as it is a multi-factor model. We will begin by describing the single-factor HJM model.

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

$$r(0,T) = T^{-1}\int_t^T f(0,t)dt \tag{7.13}$$

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 Ito'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 earlier discussion for $[0,T]$ at time t that the stochastic evolution of the forward rate can be described as:

$$df(t,T) = a(t,T)dt + \sigma(t,T)dz_t \tag{7.14}$$

or alternatively in integral form as

$$f(t,T) = f(0,T) + \int_0^t a(s,T)\mathrm{d}s + \int_0^t \sigma(s,T)dz_s \tag{7.15}$$

where a is the drift parameter, σ the volatility coefficient and z_t is the Weiner process or Brownian motion. The terms dZ or dW are sometime used to denote the Weiner process.

In (7.14) the drift and volatility coefficients are functions of time t and T. For all forward rates 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 (7.15) 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

$$r(t) = f(0,t) + \int_0^t a(s,t)ds + \int_0^t \sigma(s,t)dz_s \tag{7.16}$$

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

$$M(t) = e^{\int_0^t r(s)ds} \tag{7.17}$$

that is, the interest accumulated at the continuously compounded spot rate $r(t)$. It can be shown by substituting (7.16) into (7.17) that

$$M(t) = \exp\left[\int_0^t f(0,s)ds + \int_0^t \int_0^s a(u,s)du\, ds + \int_0^t \int_0^s \sigma(u,s)dz_u ds\right] \tag{7.18}$$

To simplify the description we write the double integrals in (7.18) in the form given below:

$$\int_0^t \int_s^t a(s,u)du\ ds + \int_0^t \int_s^t \sigma(u,s)du\ dz_s$$

The description of the process by which this simplification is achieved is relegated to a page on the website at **www.yieldcurve.com**.

Using the simplification above, it can be shown 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

$$M(t) = \exp\left[\int_0^t f(0,u)du + \int_0^d \int_s^t a(s,u)du\ ds + \int_0^t \int_s^t \sigma(s,u)du\ dz_s\right] \tag{7.19}$$

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 $B(t,T)$. 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 (7.19) is required, which is:

$$M^{-1}(t) = e^{-\int_0^t r(u)du} \tag{7.20}$$

Hence the present value at time 0 of the bond $B(t,T)$ is

$$B(t,T) = e^{-\int_0^t r(u)du} B(t,T)$$

and it can be shown that as a Weiner process the present value is given by

$$B(t,T) = \exp\left[-\int_0^T f(0,u)du - \int_0^t \int_s^T \sigma(s,u)du\ dz_s - \int_0^t \int_s^T \mathrm{a}(s,u)du\ ds\right] \tag{7.21}$$

It can be shown that the forward rate expressed as an integral is

$$f(t,T) = f(0,T) + \int_0^t a(s,T)ds + \sigma dz \tag{7.22}$$

which assumes that the forward rate is normally distributed. Crucially the different forward rates of maturity $f(0,1)$, $f(0,2)$,...,$f(0,T)$ are assumed to be perfectly correlated. The random element is the Brownian motion dz, and the impact of this process is felt over time, rather than over different maturities.

The forward rate for any maturity period T will develop as described by the drift and volatility parameters $a(t,T)$ and $\sigma(t,T)$. In the single-factor HJM model the random character of the forward-rate process is captured by the Brownian motion dz.[9] Under HJM the primary assumption is that for each T the drift and volatility processes are dependent only on the history of the Brownian motion process up to the current time t, and on the forward rates themselves up to time t.

[9]Note that certain texts use α for the drift term and W_t for the random term.

The multi-factor HJM model

Under the single-factor HJM model the movement in forward rates of all maturities is perfectly correlated. This can be too much of a restriction for market application, for example when pricing an interest-rate instrument that is dependent on the yield spread between two points on the yield curve. In the multi-factor model, each of the state variables is described by its own Brownian motion process.[10] So for example in an *m*-factor model there would be *m* Brownian motions in the model, dz_1, dz_2, ..., dz_m. This allows each *T*-maturity forward rate to be described by its own volatility level $\sigma(t,T)$ and Brownian motion process dz_i. Under this approach, the different forward rates given by the different maturity bonds that describe the current term structure evolve under more appropriate random processes, and different correlations between forward rates of differing maturities can be accommodated.

The multi-factor HJM model is given at (7.23).

$$f(t,T) = f(0,T) + \int_0^t a(s,T)ds + \sum_{i=1}^{m} \int_0^t \sigma i(s,T)dz_i(s) \tag{7.23}$$

Equation (7.23) states that the dynamics of the forward rate process, beginning with the initial rate $f(0,T)$, are specified by the set of Brownian motion processes and the drift parameter.

For practical application the evolution of the forward-rate term structure is usually carried out as a binomial-type path-dependent process. However path-independent processes have also been used. The HJM approach has become popular in the market, both for yield curve modelling and for pricing derivative instruments, due to the realistic effect of matching yield curve maturities to different volatility levels, and it is reasonably tractable when applied using the binomial-tree approach. Simulation modelling based on Monte Carlo techniques is also used. For further detail on the former approach see Jarrow (1996).

Choosing a term structure model

Selection of an appropriate term structure model is more of an art than a science. The different types of model available, and the different applications and user requirements, mean that it is not necessarily clear-cut which approach should be selected. For example a practitioner's requirements will determine whether a single-factor model or two or multi-factor model is more appropriate. The Ho–Lee and BDT models, for example, are *arbitrage* models, which means that they are designed to match the current term structure. With arbitrage (or arbitrage-free) models, assuming that the specification of the evolution of the short rate is correct, the law of no-arbitrage can be used to determine the price of interest-rate derivatives.

There are also a class of interest-rate models known as *equilibrium* models, which make an assumption of the dynamics of the short rate in the same way as arbitrage

[10] For a good introduction see chapter 6 in Baxter and Rennie (1996).

models, but are not designed to match the current term structure. With equilibrium models therefore the price of zero-coupon bonds given by the model-derived term structure is not required to (and does not) match prices seen in the market. This means that the prices of bonds and interest-rate derivatives are not given purely by the short rate process. Overall, arbitrage models take the current yield curve as described by the market prices of default-free bonds as given, whereas equilibrium models do not.

What considerations must be taken into account when deciding which term structure model to use? Adapted partly from Tuckman (1996, chapter 9) some of the key factors include:

- *Ease of application.* The key input to arbitrage models is the current spot rate term structure, which is straightforward to determine using the market price of bonds currently trading in the market. This is an advantage over equilibrium models, whose inputs are more difficult to obtain.
- *Capturing market imperfections.* The term structure generated by an arbitrage model will reflect the current market term structure, which may include pricing irregularities due to liquidity and other considerations. If this is not desired, it is a weakness of the arbitrage approach. Equilibrium models would not reflect pricing imperfections.
- *Pricing bonds and interest-rate derivatives.* Traditional seat-of-the-pants market making often employs a combination of the trader's nous, the range of prices observed in the market (often from inter-dealer broker screens) and gut feeling to price bonds. For a more scientific approach or for relative value trading[11] a yield curve model may well be desirable. In this case an equilibrium model is clearly the preferred model, as the trader will want to compare the theoretical price given by the model with the actual price observed in the market. An arbitrage model would not be appropriate because it would taken the observed yield curve, and hence the market bond price, as given, and so would assume that the market bond prices were correct. Put another way, using an arbitrage model for relative value trading would suggest to the trader that there was no gain to be made from entering into, say, a yield curve spread trade. Pricing derivative instruments such as interest-rate options or swaptions requires a different emphasis. This is because the primary consideration of the derivative market maker is the technique and price of hedging the derivative. That is, upon writing a derivative contract the market maker will simultaneously hedge the exposure using either the underlying asset or a combination of this and other derivatives such as exchange-traded futures. The derivative market maker generates profit through extracting premium and from the difference in price over time between the price of the derivative and the underlying hedge position. For this reason only an arbitrage model is appropriate, as it would price

[11] For example yield curve trades where bonds of different maturities are spread against each other, with the trader betting on the change in spread as opposed to the direction of interest rates, are a form of relative value trade.

the derivative relative to the market, which is important for a market maker; an equilibrium model would price the derivative relative to the theoretical market, which would not be appropriate since it is a market instrument that is being used as the hedge.

- *Use of models over time*. At initial use the parameters used in an interest-rate model, most notably the drift, volatility and (if applicable) mean reversion rate, reflect the state of the economy up to that point. This state is not constant, consequently over time any model must be continually recalibrated to reflect the current market state. That is, the drift rate used today when calculating the term structure may well be a different value tomorrow. This puts arbitrage models at a disadvantage, as their parameters will be changed continuously in this way. Put another way, use of arbitrage models is not consistent over time. Equilibrium model parameters are calculated from historic data or from intuitive logic, and so may not be changed as frequently. However their accuracy over time may suffer. It is up to users to decide whether they prefer the continual tweaking of the arbitrage model over the more consistent use of the equilibrium model.

This is just the beginning; there are a range of issues which must be considered by users when selecting an interest-rate model. For example, in practice it has been observed that models incorporating mean reversion work more accurately than those that do not feature this. Another factor is the computer processing power available to the user, and it is often the case that single-factor models are preferred precisely because processing is more straightforward. A good account of the different factors to be considered when assessing which model to use is given in chapter 15 of James and Webber (2000).

Appendix 7.1: Geometric Brownian motion

Brownian motion was described in 1827 by the English scientist Robert Brown, and defined mathematically by the American mathematician Norbert Weiner in 1918. As applied to the price of a security, consider the change in price of a security as it alters over time. The time now is denoted as 0, with $P(t)$ as the price of the security at time t from now. The collection of prices $P(t), 0 \leq t < \infty$ is said to follow a Brownian motion with *drift* parameter μ and variance parameter σ^2 if for all non-negative values of t and T the random variable $P(t + T) - P(t)$ is independent of all the prices P that have been recorded up to time t. That is, the historic prices do not influence the value of the random variable. Also the random variable is normally distributed with a mean μT and variance $\sigma^2 T$.

Standard Brownian motion has two drawbacks when applied to model security prices. The first and most significant is that, as the security price is a normally distributed random variable, it can assume negative values with non-negative probability, a property of the normal distribution. This cannot happen with equity prices and only very rarely, under very special conditions, with interest rates. The second drawback of standard Brownian motion is that the difference between prices over an interval is assumed to follow a normal distribution irrespective of

the price of the security at the start of the interval. This is not realistic, as the probabilities are affected by the initial price of the security.

For this reason the *geometric Brownian motion* model is used in quantitative finance. Let us consider this now. Again, the time now is 0 and the security price at time t from now is given by $P(t)$. The collection of prices $P(t), 0 \leq t < \infty$ follows a geometric Brownian motion with drift μ and standard deviation or *volatility* σ if for non-negative values of t and T the random variable $P(t + T) / P(t)$ is independent of all prices up to time t. In addition the value

$$\log\left(\frac{P(t+T)}{P(t)}\right)$$

is a normally distributed random variable with mean uT and variance $\sigma^2 T$.

What is the significance of this? It is this: once the parameters μ and σ have been ascertained, the present price of the security, and the present price only, determines the probabilities of future prices. The history of past prices has no impact. Also the probabilities of the ratio of the price at future time T to the price now are not dependent on the present price. The practical impact of this is that the probability that the price of a security doubles in price at some specified point in the future is identical whether the price now is 5 or 50.

For our purposes we need only be aware that at an initial price of $P(0)$, the expected price at time t is a function of the two parameters of geometric Brownian motion. The expected price, given an initial price $P(0)$, is given by

$$E[P(t)] = P0e^{t(\mu+\sigma^2/2)} \tag{7.24}$$

Expression (7.24) states that under geometric Brownian motion the expected price of a security is the present price increasing at the rate of $\mu + \frac{\sigma^2}{2}$.

The evolution of a price process, including an interest rate, under varying parameters is shown in Figure 7.1, with an initial price level at 100.

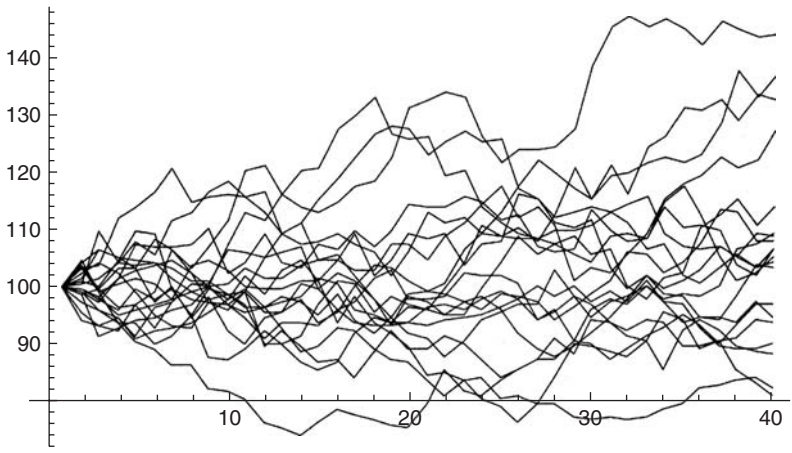


Figure 7.1 Evolution of Brownian or Weiner process

Selected references and bibliography

Baxter, M. and Rennie, A. *Financial Calculus*, Cambridge University Press 1996, pp. 57–62.
Black, F. 'Interest rates as options', *Journal of Finance*, December 1995, pp. 1371–6.
Black, F., Derman, E. and Toy, W. 'A one-factor model of interest rates and its application to treasury bond options', *Financial Analysts Journal*, Spring 1990, pp.33–9.
Brennan, M. and Schwartz, E. 'A continuous-time approach to the pricing of bonds', *Journal of Banking and Finance*, 3, 1979, pp. 133–55.
Campbell, J., 'A defence of traditional hypotheses about the term structure of interest rates', *Journal of Finance*, 41, 1986, pp.183–93.
Chen, R.-R. and Scott, L. 'Pricing interest rate options in a two-factor Cox–Ingersoll–Ross model of the term structure', *Review of Financial Studies*, Winter 1992, pp. 613–36.
Cox, J., Ingersoll, J. and Ross, S. 'A theory of the term structure of interest rates', *Econometrica*, 53, March 1985, pp. 385–407.
Fong, H. G. and Vasicek, O. 'Fixed income volatility management', *Journal of Portfolio Management*, Summer 1991, pp. 41–6.
Haug, E. *The Complete Guide to Option Pricing Formulas*, McGraw-Hill, 1998, ch. 4.
Heath, D., Jarrow, R. and Morton, A. 'Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation', *Econometrica*, 60, January 1992, pp. 77–105.
Ho, T. and Lee, S.-B. 'Term structure movements and pricing interest rate contingent claims', *Journal of Finance*, December 1986, pp.1011–29.
Hull, J. *Options, Futures and Other Derivatives*, 3rd edn, Prentice Hall, 1997, pp. 220–2.
Hull, J. and White, A. 'Pricing interest rate derivative securities', *Review of Financial Studies*, 1990, pp. 573–92.
Ingersoll, J. Jr. *Theory of Financial Decision Making*, Rowman & Littlefield, 1987.
James, J. and Webber, N. *Interest Rate Modelling*, Wiley, 2000, chs 5–15.
Jarrow, R. *Modelling Fixed Income Securities and Interest Rate Options*, McGraw-Hill, 1996.
Merton, R. *Continuous Time Finance*, Blackwell, 1993, ch. 11.
Rebonato, R. *Interest Rate Option Models*, Wiley, 1996.
Rebonato, R., Cooper, I., 'The limitations of simple two-factor interest rate models', *Journal of Financial Engineering*, March 1996.
Ross, S. M. *An Introduction to Mathematical Finance*, Cambridge University Press, 1999.
Ross, S. M. *Probability Models,* Academic Press, 2000.
Sundaresan, S. *Fixed Income Markets and their Derivatives*, South Western Publishing, 1997.
Tuckman, B. *Fixed Income Securities*, Wiley, 1996.
Van Deventer, D. and Imai, K. *Financial Risk Analytics*, Irwin, 1997, ch. 5.
Vasicek, O. 'An equilibrium characterization of the term structure', *Journal of Financial Economics*, 5, 1977, pp.177–88.

8
Fitting the Yield Curve

In this chapter we consider some of the techniques used to actually fit the term structure. In theory we could use the bootstrapping approach described earlier. For a number of reasons, however, this does not produce accurate results, and so other methods are used instead. The term structure models described in the previous chapter defined the interest rate process under various assumptions about the nature of the stochastic process that drives these rates. However the zero-coupon curve derived by models such as those described by Vasicek (1977), Brennan and Schwartz (1979) and Cox, Ingersoll and Ross (1985) do not fit the observed market rates or spot rates implied by market yields, and generally market yield curves are found to contain more variable shapes than those derived using term structure models. Hence the interest rate models described in Chapter 7 are required to be *calibrated* to the market, and in practice they are calibrated to the market yield curve. This is carried out in two ways: either the model is calibrated to market instruments such as money market products and interest-rate swaps, which are used to construct the yield curve, or the yield curve is constructed from market instrument rates and the model is calibrated to this constructed curve. If the latter approach is preferred, there are a number of non-parametric methods that may be used. We consider these later.

In this chapter we present an overview of some of the methods used to fit the yield curve. For the interested reader a selection of useful references is given in the bibliography.

Yield curve smoothing[1]

Introduction

An approach that has been used to estimate the term structure was described by Carleton and Cooper (1976). It assumed that default-free bond cash flows are payable on specified discrete dates, with a set of unrelated discount factors that apply to each cash flow. These discount factors were then estimated as regression coefficients, with each bond cash flow acting as the independent variables, and

[1]Large parts of this section previously appeared in Choudhry (2001).

the bond price for that date acting as the dependent variable.[2] Using simple linear regression in this way produces a discrete discount function, not a continuous one, and forward rates that are estimated from this function are very jagged.

An approach more readily accepted by the market was described by McCulloch (1971), who fitted the discount function using polynomial splines. This method produces a continuous function, and one that is linear so that the *ordinary least squares* regression technique can be employed. In a later study, Langetieg (1981)[3] uses an extended McCulloch method, fitting *cubic splines* to zero-coupon rates instead of the discount function, and using non-linear methods of estimation.

That is the historical summary of early efforts, but let us get back to the beginning. We know that the term structure can be described as the complete set of discount factors, the discount function, which can be extracted from the price of default-free bonds trading in the market. The bootstrapping technique described in Chapter 4 may be used to extract the relevant discount factors. However there are a number of reasons that this approach is problematic in practice. First, it is unlikely that the complete set of bonds in the market will pay cash flows at precise six-month intervals every six months from today to 30 years or longer. An adjustment is made for cash flows received at irregular intervals, and for the lack of cash flows available at longer maturities. Another issue is the fact that the technique presented earlier allowed practitioners to calculate the discount factor for six-month maturities, whereas it may be necessary to determine the discount factor for non-standard periods, such as four-month or 14.2-year maturities. This is often the case when pricing derivative instruments.

A third issue concerns the market price of bonds. These often reflect specific investor considerations, which include:

- the liquidity or lack thereof of certain bonds, caused by issue sizes, market maker support, investor demand, non-standard maturity and a host of other factors
- the fact that bonds do not trade continuously, so that some bond prices will be 'newer' than others
- the tax treatment of bond cash flows, and the effect that this has on bond prices
- the effect of the bid–offer spread on the market prices used.

The statistical term used for bond prices subject to these considerations is *error*. It is also common to come across the statement that these effects introduce *noise* into market prices.

Smoothing techniques

A common technique that may be used, but which is not accurate enough and so is not recommended for market use, is *linear interpolation*. In this approach the

[2] The basics of regression are summarised in Appendix 5.1 of Choudhry (2001). Readers who wish to get a firm grasp of econometric techniques used in financial market analysis should consult Gujarati (1995).

[3] Reference in Vasicek and Fong (1982).

set of bond prices are used to graph a redemption yield curve (as in the previous section), and where bonds are not available for the required maturity term, the yield is interpolated from actual yields. Using UK gilt yields for 9 November 2000 we plot this as shown in Figure 8.1. The interpolated yields are marked on the x axis. Figure 8.1 looks reasonable for any practitioner's purpose. However spot and forward yields that are obtained from this curve using the linear interpolation technique are apt to behave in unrealistic fashion, as shown in Figure 8.2.[4] The forward curve is very bumpy, and each bump will correspond to a bond used in the original set. The spot rate has a kink at three years and again at nine years, and so the forward curve jumps significantly at these points. This curve would appear to be particularly unrealistic.

For this reason, market analysts do not usually consider linear interpolation and instead use exponential interpolation, multiple regression or spline-based methods. One approach might be to assume a functional form for the discount function and estimate parameters of this form from the prices of bonds in the market. We consider these approaches next.

Using a cubic polynomial

A simple functional form for the discount function is a *cubic polynomial*. This approach consists of approximating the set of discount factors using a cubic

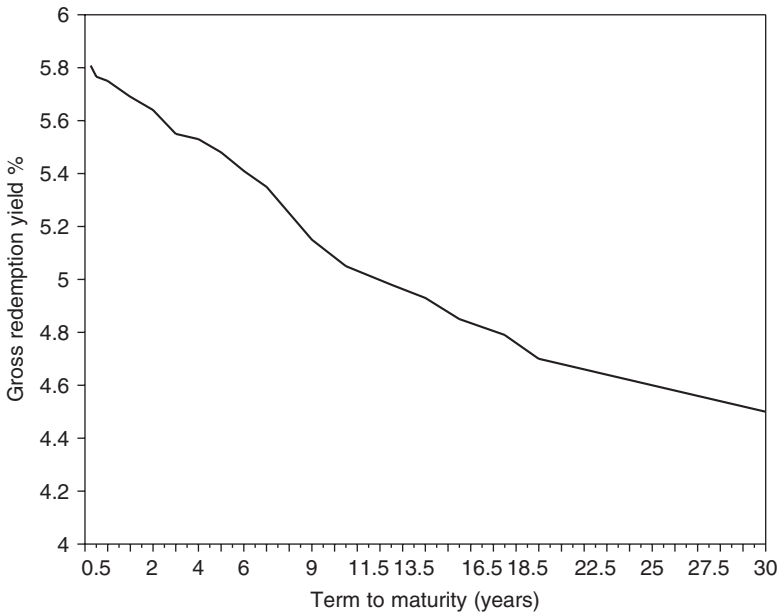


Figure 8.1 Linear interpolation of bond yields, 9 November 2000
Source: Bloomberg.

[4] The spot and forward yield curves were calculated using the RATE application software. This can be downloaded from www.yieldcurve.com

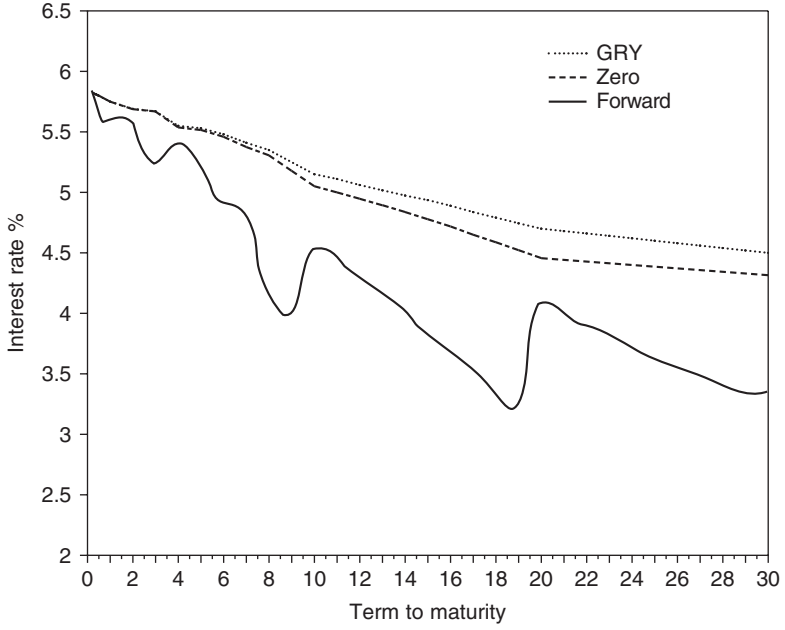


Figure 8.2 Spot and forward rates implied from rates in Figure 8.1

function of time. If we say that $d(t)$ is the discount factor for maturity t, we approximate the set of discount factors using the following cubic function.

$$\hat{d}(t) = a_0 + a_1(t) + a_2(t)^2 + a_3(t)^3 \tag{8.1}$$

In some texts the coefficients sometimes are written as *a, b*, and *c* rather than a_1 and so on.

The discount factor for $t = 0$, that is at time now, is 1. Therefore $a_0 = 1$, and (8.1) can then be rewritten as:

$$\hat{d}(t) - 1 = a_1(t) + a_2(t)^2 + a_3(t)^3 \tag{8.2}$$

The market price of a traded coupon bond can be expressed in terms of discount factors. So at (8.3) we show the expression for the price of an *N*-maturity bond paying identical coupons *C* at regular intervals and redeemed at maturity at *M*.

$$P = d(t_1)C + d(t_2)C + ... + d(t_N)(C + M) \tag{8.3}$$

Using the cubic polynomial equation (8.2), expression (8.3) is transformed into:

$$P = C[1 + a_1(t_1) + a_2(t_1)^2 + a_3(t_1)^3 + ... + (C + M)[1 + a_1(t_N) + a_2(t_N)^2 + a_3(t_N)^3] \tag{8.4}$$

We require the coefficients of the cubic function in order to start describing the yield curve, so we rearrange (8.4) in order to express it in terms of these coefficients. This is shown at (8.5).

$$P = M + \Sigma C + a_1[C(t_1) + ... + (C + M)(t_N)] + a_2[C(t_1)^2 + ... + (C + M)(t_N)^2]$$
$$+ a_3[C(t_1)^3 + ... + (C + M)(t_N)^3] \tag{8.5}$$

In the same way we can express the pricing equation for each bond in our data set in terms of the unknown parameters of the cubic function.

From (8.5) we may write:

$$P = (M + \Sigma C) = a_1 X_1 + a_2 X_2 + a_3 X_3 \tag{8.6}$$

where X_i is the appropriate expression in square brackets in (8.5); this is the form in which the expression is commonly encountered in text books.

In practice the cubic polynomial approach is too limited a technique, requiring one equation per bond, and does not have the required flexibility to fit market data satisfactorily. The resulting curve is not really a curve but rather a set of independent discount factors that have been fitted with a line of best fit. In addition the impact of small changes in the data can be significant at the non-local level, so for example a change in a single data point at the early maturities can result in badly behaved longer maturities. Alternatively a *piecewise cubic polynomial* approach is used, whereby $d(t)$ is assumed to be a different cubic polynomial over each maturity range. This means that the parameters a_1, a_2 and a_3 will be different over each maturity range. We will look at a special case of this use, the cubic spline, a little later.

Non-parametric methods

Besides the cubic polynomial approach described in the previous section there are two main approaches to fitting the term structure. These are usually grouped into *parametric* and *non-parametric* curves. Parametric curves are based on term structure models such as the Vasicek model or Longstaff and Schwartz model. Non-parametric curves are not derived from an interest-rate model and are general approaches, described using a set of parameters. They include spline-based methods.

Spline-based methods

A spline is a statistical technique and a form of linear interpolation method. There is more than one way of applying splines, and the most straightforward method for understanding the process is the spline function fitted using regression techniques. For the purposes of yield curve construction this method can cause curves to jump wildly, and it is over-sensitive to changes in parameters.[5]

[5] For instance see James and Webber (2000), section 15.3.

However we feel it is the most accessible method. An introduction to the basic technique, as described in Suits, Mason and Chan (1978), is given in Appendix 5.2 of Choudhry (2001).[6]

An *n*-th order spline is a piecewise polynomial approximation with *n*-degree polynomials that are differentiable *n*–1 times. *Piecewise* means that the different polynomials are connected at arbitrarily selected points known as *knot points* (see Appendix 5.2 of Choudhry 2001). A cubic spline is a three-order spline, and is a piecewise cubic polynomial that is differentiable twice along all its points.

The x axis in the regression is divided into segments at arbitrary points known as knot points. At each knot point the slope of adjoining curves is required to match, as must the curvature. Figure 8.3 is a cubic spline. The knot points are selected at 0, 2, 5, 10 and 25 years. At each of these points the curve is a cubic polynomial, and with this function we could accommodate a high and low in each space bounded by the knot points.

Cubic spline interpolation assumes that there is a cubic polynomial that can estimate the yield curve at each maturity gap. One can think of a spline as a number of separate polynomials of $y = f(X)$ where X is the complete range, divided into user-specified segments that are joined smoothly at the knot points. If we have a

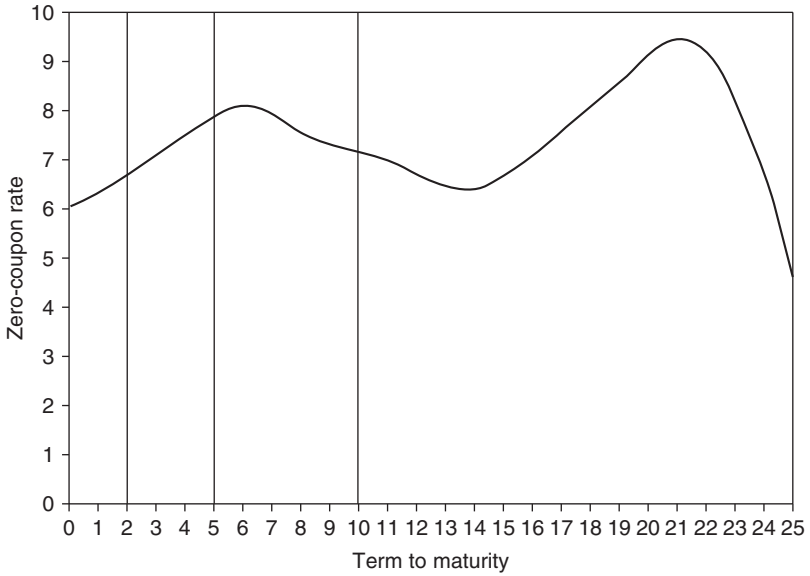


Figure 8.3 Cubic spline with knot points at 0, 2, 5, 10 and 25 years

[6]The original article by Suits, Mason and Chan (1978) is excellent and highly recommended.

set of bond yields $r_0,r_1,r_2,\ldots,r_n$ at maturity points $t_0,t_1,t_2,\ldots,t_n$, we can estimate the cubic spline function in the following way:

- The yield on bond i at time t is expressed as a cubic polynomial of the form $r_i(t) = a_i + b_it + c_it^2 + d_it^3$ for the interval over t_i and t_i^{-1}.
- The coefficients of the cubic polynomial are calculated for all n intervals between the $n + 1$ data points, which results in $4n$ unknown coefficients that must be computed.
- These equations can be solved because they are made to fit the observed data. They are twice differentiable at the knot points, and these derivatives are equal at these points.
- The constraints specified are that the curve is instantaneously straight at the start of the curve (the shortest maturity) and instantaneously straight at the end of the curve, the longest maturity, that is, $r^n(0) = 0$.

An accessible and readable account of this technique can be found in Van Deventer and Imai (1997).

The general formula for a cubic spline is:

$$s(\tau) = \sum_{i=0}^{3} a_i \tau^i + \frac{1}{3!} \sum_{p=1}^{n-1} bp(\tau - X_p)^3 \tag{8.7}$$

where τ is the time of receipt of cash flows and X_p refers to the points where adjacent polynomials are joined, which are known as knot points, with $\{X_0,\ldots,X_n\}$, $X_p < X_{p+1}$, $p = 0,\ldots,n-1$. In addition, $(\tau - X_p) = \max(\tau - X_p,0)$. The cubic spline is twice differentiable at the knot points. In practice the spline is written down as a set of basis functions, with the general spline being made up of a combination of these. One way to do this is by using what are known as *B-splines*. For a specified number of knot points $\{X_0,\ldots,X_n\}$ this is given by (8.8):

$$B_p(\tau) = \sum_{j=p}^{p+4} \left(\prod_{i=p,1\neq i}^{p+4} \frac{1}{Xi - Xj} \right) (\tau - X_p)^3 \tag{8.8}$$

where $B_p(\tau)$ are cubic splines which are approximated on $\{X_0,\ldots,X_n\}$ with the following function:

$$\delta(\tau) = \delta(\tau | \lambda_{-3},\ldots,\ \lambda_{n-1}) = \sum_{p=-3}^{n-1} \lambda_p B_p(\tau) \tag{8.9}$$

with $\lambda = \{\lambda_{-3},\ldots,\ \lambda_{n-1}\}$ the required coefficients. The maturity periods $\tau_1,\ldots,\tau_n$ specify the B-splines so that $B = \{B_p(\tau_j)\}_{p\,=\,3,\ldots,n-1,\ j\,=\,1,\ldots m}$ and $\delta = (\delta(\tau_1),\ldots,\delta(\tau_m))$. This allows us to set

$$\delta = B'\lambda \tag{8.10}$$

and therefore the regression equation

$$\lambda^* = \arg\min_{\lambda} \{\varepsilon'\varepsilon \mid \varepsilon = P - D\lambda\} \tag{8.11}$$

with $D = CB'$.

$\varepsilon'\varepsilon$ are the minimum errors. The regression at (8.11) is computed using ordinary least squares regression.

An advanced illustration of the use of B-splines is given in the next chapter. Appendix 5.2 of Choudhry (2001) provides background on splines fitted using regression methods.

Nelson and Siegel curves

The curve fitting technique first described by Nelson and Siegel (1985) has since been applied and modified by other authors, which is why they are sometimes described as a 'family' of curves. These curves provide a satisfactory rough fit of the complete term structure, with some loss of accuracy at the very short and very long end. In the original curve the authors specify four parameters. The approach is not a bootstrapping technique, rather a method for estimating the zero-coupon rate function from the yields observed on T-bills, under an assumed function for forward rates.

The Nelson and Siegel curve states that the implied forward rate yield curve may be modelled along the entire term structure using the following function:

$$rf(m,\beta) = \beta_0 + \beta_1 \exp\left(\frac{-m}{t_1}\right) + \beta_2 \left(\frac{m}{t_1}\right) \exp\left(\frac{-m}{t_1}\right) \tag{8.12}$$

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 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.

A version of the Nelson and Siegel curve is the Svensson model (1994) with an adjustment to allow for the humped characteristic of the yield curve. This is fitted by adding an extension, as shown by (8.13).

$$rf(m,\beta) = \beta_0 + \beta_1 \exp\left(\frac{-m}{t_1}\right) + \beta_2 \left(\frac{m}{t_1}\right) \exp\left(\frac{-m}{t_1}\right) + \beta_3 \left(\frac{m}{t_2}\right) \exp\left(\frac{-m}{t_2}\right) \tag{8.13}$$

The Svensson curve is modelled therefore using six parameters, with additional input of β_3 and t_2.

Nelson and Siegel curves are popular in the market because they are straightforward to calculate. Jordan and Mansi (2000) state that one of the advantages of these curves is that they force the long-date forward curve into an horizontal asymptote, while another is that the user is not required to specify knot points, the

choice of which determines the effectiveness or otherwise of cubic spline curves. The disadvantage they note is that these curves are less flexible than spline-based curves and there is therefore a chance that they will not fit the observed data as accurately as spline models.[7] James and Webber (2000, pp. 444–5) also suggest that Nelson and Siegel curves are slightly inflexible due to the limited number of parameters, and are accurate for yield curves that have only one hump, but are unsatisfactory for curves that possess both a hump and a trough. As they are only reasonable for approximations, Nelson and Siegel curves would not be appropriate for no-arbitrage applications.

Comparing curves

Whichever curve is chosen will depend on the user's requirements and the purpose for which the model is required. The choice of modelling methodology is usually a trade-off between simplicity and ease of computation and accuracy. Essentially the curve chosen must fulfil the qualities of:

- Accuracy: is the curve a reasonable fit of the market curve? Is it flexible enough to accommodate a variety of yield curve shapes?
- Model consistency: is the curve fitting method consistent with a theoretical yield curve model such as Vasicek or Cox–Ingersoll–Ross?
- Simplicity: is the curve reasonably straightforward to compute? That is, is it tractable?

The different methodologies all fit these requirements to a greater or lesser extent. A good summary of the advantages and disadvantages of some popular modelling methods can be found in James and Webber (2000, ch. 15).

Selected bibliography and references

Anderson, N., Breedon, F., Deacon, M., Derry, A. and Murphy, M. *Estimating and Interpreting the Yield Curve*, Wiley 1996.

Brennan, M. and Schwartz, E. 'A continuous-time approach to the pricing of bonds', *Journal of Banking and Finance*, 3, 1979, pp. 133–55.

Carleton, W. and Cooper, I. 'Estimation and uses of the term structure of interest rates', *Journal of Finance*, September 1976, pp. 1067–83.

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

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

Deacon, M. and Derry, A. 'Estimating the term structure of interest rates', Bank of England Working Paper no. 24, July 1994.

Gujarati, D. *Basic Econometrics*, 3rd edn, McGraw-Hill, 1995.

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

[7] This is an excellent article, strongly recommended. A good overview introduction to curve fitting is given in the introduction, and the main body of the article gives a good insight into the type of research that is currently being undertaken in yield curve analysis.

Jordan, J. and Mansi, S. 'How well do constant-maturity treasuries approximate the on-the-run term structure?', *Journal of Fixed Income*, 10, 2, September 2000, pp. 35–45.
Langetieg, T. C. 'A multivariate model of the term structure', *Journal of Finance*, 35, 1980, pp. 71–91.
McCulloch, J. H. 'Measuring the term structure of interest rates', *Journal of Business*, January 1971, pp. 19–31.
Nelson, C. and Siegel, A. 'Parsimonious modelling of yield curves', *Journal of Business*, 6, 1987, pp. 473–83.
Questa, G. *Fixed Income Analysis for the Global Financial Market*, Wiley, 1999.
Steeley, J. M. 'Estimating the gilt-edged term structure: basis splines and confidence intervals', *Journal of Business Finance and Accounting*, 18, 1991, pp. 513–30.
Suits, D., Mason, A. and Chan, L. 'Spline functions fitted by standard regression methods', *Review of Economics and Statistics*, 60, 1978, pp. 132–9.
Tuckman, B. *Fixed Income Securities*, Wiley, 1996.
Van Deventer, D. and Imai, K. *Financial Risk Analytics*, Irwin, 1997.
Vasicek, O. 'An equilibrium characterization of the term structure', *Journal of Financial Economics*, 5, 1977, pp. 177–88.
Vasicek, O. and Fong, H. G. 'Term structure modelling using exponential splines', *Journal of Finance*, 37, 2, May 1982, pp. 339–48.

9

B-Spline Modelling and Fitting the Term Structure

Introduction

For market practitioners, zero-coupon rate curves are the basic tools used to value interest-rate based instruments. Curves are built using market data such as money market rates, swap rates, interest rates futures or bond prices as inputs. Despite the name, it is not in fact the 'zero coupon' rates that are the most important output from a curve fitting methodology, but rather a set of quantities known as *discount factors*. It is these that are crucial for the pricing of interest rate-based instruments.

In this chapter, we provide an advanced methodology to extract discount factors from a set of bond prices. The objective is to be as explicit as possible so that non-mathematicians may be able to incorporate the methodology into their daily activity.

We begin with basic definitions; more experienced readers can skip this and directly go to page 188.

Basic concepts

A *zero-coupon rate* is the interest rate that is generated by an investment in cash over a certain period of time. The name comes from the fact that no intermediate payment is made to the investor, and there is only the one cash flow on maturity. Zero-coupon rates are usually expressed on an annual basis. A zero-coupon rate is fully described by its *value* (such as 8% or 10%), its period (for example, two years), its day count convention (for example 30/36 or actual/365) and its compounding frequency (such as annual, semi-annual and so on). Day count conventions are considered elsewhere in this book. We consider two simple examples to illustrate the basic concept.

Example 9.1

Consider a two-year zero coupon rate, on 30/360, annual basis, that is, 10% for value on 1 January 2001. If an investor were to invest \$100 on

01/01/2001 until 31/12/2002 he would get £121 at the end of the two-year period, given by

$$121 = 100[investment] \times (1 + 10\%)[year1] \times (1 + 10\%)[year2]$$

In this case the day count convention is irrelevant as the period of the zero coupon is an exact number of years.

Example 9.2

In this example we see how the day count convention is relevant. Assume this same investor were to invest £100 on 1 January 2001 until 14 June 2002 and that the corresponding zero coupon rate for the period is again 10% (30/360, annual). The investor would receive \$114.88 at the end of the period, given by:

$$114.88 = 100 \times (1 + 10\%) \times (1 + 10\%)^{(150 + 14)/360}$$

This reflects the fact that in 2002 there will be five full months (i.e. 5×30 or 150 days) plus 14 days in June, for a total of 164 days.

Present value and discount factor

The *present value* of a future cash flow is its value today. For instance using the data in Example 9.1, the present value of a two-year cash flow of 121 is 100. A *discount factor* is defined by a *date* and an *amount*. It is the coefficient by which we need to multiply a future cash flow to obtain its present value.

A discount factor is always less or equal to 1 and tends to zero when its date tends to infinity.

Again in Example 9.1 the amount of the two-year discount factor is 100/121 or 0.8264463.

Bootstrapping

We have already described in Chapter 4 the bootstrapping methodology to extract a zero-coupon curve from a set of bonds. This methodology offers some comfort and peace of mind, because when recalculating the price of the bonds used in the application algorithm, we back out their original market value. That is, the technique allows us to construct zero rates that fit the original bond market prices. This 'perfect match' allows for easy testing and benchmarking of the implementation of the method. Nevertheless, it is far from being satisfactory: because of some major distortions in the resulting zero-curve computed, it often leads to negative forward rates and to unrealistic figures as far as the zero-coupon rates themselves are concerned.

For this reason market practitioners require other more advanced techniques, to get around the unsuitable results obtained from the traditional technique. This more advanced analytic is used by traders to price interest-rate instruments, as well as by risk managers to measure the risk exposure entered into and the risk-adjusted profit performance of the front book.

An advanced methodology: the cubic B-spline

The methodology we describe now is based on a *cubic B-spline* representation of the discount curve. The reasoning behind this approach is that B-splines provide very smooth curves that are very easy to manipulate if necessary; for instance it is relatively straightforward to make a B-spline curve go through specific points. (B-splines are described later in this chapter.)

The methodology uses a least squares method to try to minimise the gap between a set of bonds' theoretical prices, that is, the prices that would be obtained using the advanced methodology, and their observed market prices. This is illustrated with an example that was built using UK government bonds or gilts. Of course it is perfectly permissible to use, say, AAA-rated corporate bonds from various institutions, to obtain a discount curve for AAA rated companies.

In their 1982 paper, Vasicek, Oldrich and Fong offered an alternative method using 'third order exponential splines'. This essentially assumes that discount curves are a linear combination of exponential functions within given time buckets. Alternatively, it can be seen as a more complex method than the one described in this book: instead of using splines to describe the discount curve itself, it operates a change of variable in the discount factor function from t (the time) to the new variable x, using the logarithm function.

$$t = -\frac{1}{\alpha}\log(1-x)$$

The corresponding new discount function is then approximated using splines.

Finally, they solve a minimisation problem where α (called 'the limiting value of the forward rates') is the unknown and where the function to minimise is itself a minimisation problem. However this approach is not fully described here; rather it can be seen as a comparison with or an upgrade from our proposed methodology. Once the basic concepts of splines and optimisation (least square methods) are well understood and have been implemented, it is possible to switch methodology with a limited amount of effort.

Notation

We describe the methodology using the following notation:

N is the number of bonds used to compute the discount curve.

Q is the S-dimensional control vector whose components $(Q_1)_{I=0..S-1}$ are the unknowns of the optimisation problem.

P_i^{th} is the theoretical value of bond number *i* computed using the advanced methodology.
P_i^{ma} is the market value (dirty price) of bond number *i*.
n_i is the number of cash flows for bond number *i*.
C^i_j is the cash flow number *j* of bond number *i*.
x^i_j is the date at which C^i_j is expected to be paid.
$Df(x)$ is the discount factor computed using the advanced methodology for a given date *x*.
$B_{k,3}$ is the cubic B-spline number *k*.
$(t_i)_{I=0..S+3}$ is a set of dates used to define the B-splines.

By definition, for the discount factor we have

$$Df(x) = \sum_{k=0}^{S-1} Q_k B_{k,3}(x) \tag{9.1}$$

Using the above notation we can express the theoretical value of a bond, that is its theoretical dirty price as

$$P_i^{th} = \sum_{j=1}^{n_i} C_j^i Df(x_j^i) = \sum_{j=1}^{n_i} C_j^i \sum_{k=0}^{S-1} Q_k B_{k,3}(x_j^i) \tag{9.2}$$

Our objective is to minimise the expression

$$F(Q) = \sum_{i=1}^{N} (P_l^{th} - P_l^{ma})^2 \tag{9.3}$$

We have

$$\begin{aligned} Min\ F(Q) \quad & \Leftrightarrow \underset{Q\in\Re^s}{Min} \sum_{i=1}^{N} (P_i^{th} - P_i^{ma})^2 \\ & \Leftrightarrow \underset{Q\in\Re^s}{Min} \sum_{i=1}^{N} \left(\sum_{j=1}^{n_i} C_j^i \sum_{k=0}^{S-1} Q_k B_{k,3}(x_j^i) - P_i^{ma} \right)^2 \\ & \Leftrightarrow \underset{Q\in\Re^s}{Min} \sum_{i=1}^{N} \left(\sum_{k=0}^{S-1} Q_k \sum_{j=1}^{n_i} C_j^i B_{k,3}(x_j^i) - P_i^{ma} \right)^2 \end{aligned}$$

By defining

$$a_k^i = \sum_{j=1}^{n_i} C_j^i B_{k,3}(x_j^i) \tag{9.4}$$

the minimisation problem can also be written as

$$\underset{Q\in\Re^s}{Min} \sum_{i=1}^{N} \left(\sum_{k=0}^{S-1} Q_k a_k^i - P_i^{ma} \right)^2 \tag{9.5}$$

This is equivalent to looking for Q such that $\nabla F(Q) = \bar{0}$.

We express every component l of $\nabla F(Q)$ as

$$(\nabla F(Q))_l = \partial_{Q1} F(Q) \tag{9.6}$$

where

$$\partial_{Q1} F(Q) = 2\sum_{i=1}^{N}\left(\sum_{k=0}^{S-1} Q_k a_k^i - P_i^{ma}\right) a_j^i \tag{9.7}$$

We are therefore looking for Q such that

$$\sum_{i=1}^{N}\sum_{k=0}^{S-1} Q_k a_k^i a_l^i = \sum_{i=1}^{N} P_i^{ma}\ a_l^i \tag{9.8}$$

This can also be written as

$$\sum_{k=0}^{S-1} Q_k \sum_{i=1}^{N} a_k^i a_l^i = \sum_{i=1}^{N} P_i^{ma} a_l^i \tag{9.9}$$

In conclusion then, the minimisation problem we wish to solve is equivalent to solving the $S \times S$ linear system

$$AQ = B \tag{9.10}$$

where

$$(A_{l,k})_{\substack{l=0\ldots S-1 \\ k=0\ldots S-1}} = \sum_{i=1}^{N} a_k^i\ a_l^i$$

and

$$(B_l)_{l=0\ldots S-1} = \sum\nolimits_{i=1}^{N} P_i^{ma} a_l^i$$

(A) is obviously symmetric, which will allow the use of specific algorithms to solve the linear system of equations. For this illustration however we use the GMRES algorithm, described below.

Simplification

If all the bonds i under consideration pay a fixed coupon C^i and are redeemed at par, we have

$$a_k^i = C^i \sum_{j=1}^{n_i=1} B_{k,3}(x_j^i) + 100\ B_{n_i}(x_{n_i}^i)$$

This assumes that prices are expressed as percentages of par.

A priori, the minimisation problem described above is constraint-free. However we wish to introduce the following constraints, as we know of two properties for discount curves that we wish to enforce in our methodology:

- the discount factor for today is 1
- the discount factor curve tends to zero when time tends to infinity.

Ideally we want to have to deal with a constraint-free minimisation problem, so we embed these two constraints in the definition of $F(Q)$ itself. This is carried out as follows.

For the constraint that *Df* for today is 1 we set:

$$\sum_{k=0}^{S-1} Q_k B_{k,3}\ (Today) = 1 \tag{9.11}$$

By choosing the four first t_i equal to today ($t_0 = t_1 = t_2 = t_3 = Today$) and using properties of the B-splines (discussed elsewhere), we only have one non-zero term in the above sum, which leaves us with

$$Q_0 = 1 \tag{9.12}$$

Thus Q_0 is no longer an unknown in the problem; the linear system to solve becomes an $S - 1 \times S - 1$ system. Therefore the matrix A needs to be amended accordingly; the first row and first column disappear and the second member of the linear system becomes B' with

$$\forall k = 1 \ldots S-1\ B'_k = B_k - A_{k,0}$$

For the second constraint, by choosing t_{s+3} equal to a 'very big' number (in our example we chose 100,000 to 'approximate' to infinity), we ensure that the discount curve smoothly goes down to zero for this number. This is a satisfactory estimate for the purpose of having the discount curve going down to zero when time tends to infinity.

Example 9.3: Fitting the gilt zero-coupon yield curve

In this example, we compute a zero curve on 30 June 1997 from UK gilt prices.[1] In it we use

- 30 different bonds, so that N = 30
- 21 dates to define the B-splines, so that S = 17.

[1] Note that at that time gilt prices were quoted in 'ticks' or 32nds, similar to US Treasury price quotes.

Table 9.1 t_i values

t_0	30 June 1997	t_{11}	07 June 2006	t_{14}	07 June 2010
t_1	30 June 1997	t_{12}	07 Aug 2007	t_{15}	07 June 2013
t_2	30 June 1997	t_{13}	07 Dec 2008	t_{16}	07 June 2015
t_3	30 June 1997	t_7	07 June 2000	t_{17}	25 Aug 2017
t_4	07 Sept 1997	t_8	07 Dec 2001	t_{18}	25 Aug 2019
t_5	30 March 1998	t_9	07 June 2003	t_{19}	08 June 2021
t_6	07 June 1999	t_{10}	07 Dec 2004	t_{20}	100000

In Table 9.1 we show our choices for the t_i values. Note the following:

- On a daily basis, once the required bonds and the set of t_i are known, the matrix only needs to be computed once. As changes in market prices occur, the discount curve can be updated quickly by only computing again the second member of the linear system and then solving this.
- The methodology can be enhanced to take into account the fact that some bonds are benchmarks whereas others are very illiquid. Investors usually ask for a yield premium to hold less liquid bonds. The direct effect when computing a discount curve is that it tends to distort the expected results. To cater for this, the original function can be changed by incorporating weights (w_i) in the model. $F(Q)$ then becomes

$$F(Q) = \sum_{i=1}^{N} w_i (P_i^{th} - P_i^{ma})^2$$

  The idea here is that when a bond is a benchmark issue, its corresponding weight is going to be high, whereas for a very illiquid bond, the weight is going to be low.
- The choice of the t_i values will influence the accuracy of the model. Apart from the first four values and the last value, they must be connected to the maturity of the bonds used in the calculation. The more maturities we have in a given period, the more t_i we have to choose in that period. Of course there is no need to choose many t_i values in a period when no bonds mature.

Table 9.2 shows the bonds that were used, together with their market price and the corresponding theoretical price for each bond, which has been computed using our methodology. The difference between the two prices is given in the column headed 'pence spread'.

The discount curve obtained and the corresponding zero-curve are shown in Figure 9.1. For validation purposes, we have also included the three-month forward curve.

Table 9.2 UK gilt observed market and theoretical prices

27 Jun 1997 for settlement 30 June 1997

Coupon	*Maturity*	*Market price*	*Accrued*	*GRY*	*Theo. price*	*Theo. GRY*	*bp spread*	*pence spread*
8.75%	01 Sept 1997	100–11	2.901	6.411	100.315	6.576	−16.5	3
9.75%	19 Jan 1998	101–21	4.327	6.705	101.704	6.616	8.8	−5
7.25%	30 Mar 1998	100–09	1.827	6.798	100.218	6.885	−8.8	6
12.25%	26 Mar 1999	108–14	3.222	6.972	108.486	6.944	2.8	−5
6.00%	10 Aug 1999	98–08	2.301	6.913	98.163	6.958	−4.6	9
9.00%	03 Mar 2000	104–20	2.934	7.052	104.649	7.042	1.0	−2
13.00%	14 July 2000	115–28	5.948	7.115	115.941	7.092	2.3	−7
8.00%	07 Dec 2000	102–27	0.504	7.048	102.685	7.100	−5.2	16
10.00%	26 Feb 2001	109–04	3.397	7.128	109.211	7.102	2.6	−9
7.00%	06 Nov 2001	99–22	1.055	7.073	99.690	7.073	0.1	0
7.00%	07 June 2002	99–27	0.441	7.034	99.652	7.081	−4.7	19
9.75%	27 Aug 2002	111–04	3.286	7.139	111.346	7.091	4.8	−22
8.00%	10 June 2003	104–10	0.438	7.095	104.249	7.108	−1.3	6
9.50%	25 Oct 2004	113–14	1.718	7.108	113.528	7.093	1.5	−9
6.75%	26 Nov 2004	98–05	0.647	7.068	98.006	7.094	−2.7	15
9.50%	18 Apr 2005	114–02	1.900	7.115	114.167	7.098	1.6	−10
8.50%	07 Dec 2005	108–23	0.536	7.105	108.654	7.115	−1.0	6
7.75%	08 Sept 2006	104–04	2.421	7.124	104.155	7.119	0.4	−3
7.50%	07 Dec 2006	102–21	0.473	7.106	102.594	7.115	−0.9	6
8.50%	16 Jul 2007	109–22	3.842	7.137	109.893	7.109	2.7	−21
7.25%	07 Dec 2007	101–06	0.457	7.085	100.995	7.111	−2.6	19
9.00%	13 Oct 2008	114–08	1.923	7.136	114.271	7.134	0.2	−2
8.00%	25 Sept 2009	106–24	2.126	7.157	106.775	7.154	0.3	−2
6.25%	25 Nov 2010	92–02	0.616	7.178	92.095	7.174	0.4	−3
9.00%	12 July 2011	116–01	4.167	7.173	115.987	7.177	−0.5	4
9.00%	06 Aug 2012	116–19	3.551	7.184	116.574	7.186	−0.2	2
8.00%	27 Sept 2013	107–24	2.082	7.179	107.786	7.175	0.4	−4
8.00%	07 Dec 2015	108–23	0.504	7.140	108.702	7.142	−0.2	2
8.75%	25 Aug 2017	116–18	2.997	7.184	116.569	7.183	0.1	−1
8.00%	07 June 2021	109–31	0.504	7.125	109.968	7.125	0.0	0

Mathematical tools

In this chapter, we have used a pseudo iterative method to solve a linear problem, known as GMRES, and a B-spline interpolation method for the discount factor curve. We now describe these two mathematical tools.

GMRES

GMRES is a methodology used to solve multi-dimensional linear problems; the acronym comes from General Minimum RESidual. It has been used heavily in aeronautical engineering and aerodynamics since the early 1990s, because it has proven itself to be a very efficient solver. We have direct experience of using the method while at Dassault Aviation, to model flow behaviour around an airfoil and a rocket. Both applications were non linear problems using the two-dimensional compressible potential equations and the two-dimensional Euler equations.

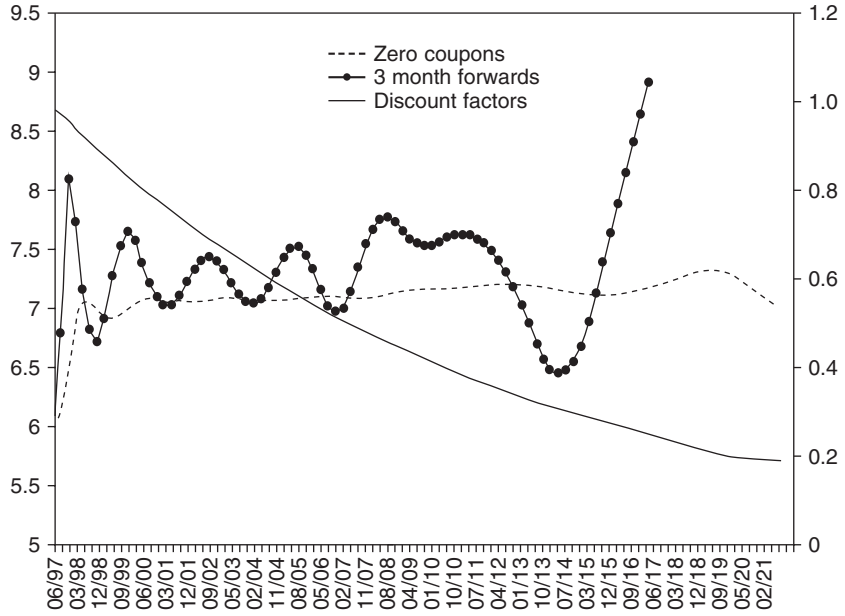


Figure 9.1 Calculated discount function, spot and forward curves

Remember that to solve a linear problem $Ax = b$, where A is a n $\times$ n matrix, x is the unknown n-vector and b is the second-member n-vector, there are two families of algorithm we can use: direct methods and iterative methods.

Direct methods include the intuitive one readers will have learnt at school, when one has to invert the matrix A to obtain A^{-1}. The solution of the problem can then be directly obtained, by multiplying b by A^{-1}, that is:

$$x = A^{-1}b$$

Direct methods have one main drawback: when the dimension of the problem increases, so that n goes up, the number of operations required to invert the matrix A goes up dramatically, as well as the computer memory required to store all the necessary data. Therefore in some cases the calculation cannot realistically be performed.

For this reason mathematicians have introduced iterative methods. These consist of starting from an estimation of the solution, performing some operations to obtain a more precise solution and then performing the same kind of operation over and over again, until a predefined error is reached.

For instance, assume the initial estimation of the solution is x_0. The initial error will then be

$$\| b - Ax_0 \|$$

where $\| \bullet \|$ is the norm of the vector $b - Ax_0$.

Then the following error will be $\|b - Ax_1\|$ where x_1 is the second estimation calculated, and so on.

The iterative algorithm will then certainly lead us to the solution when

$$\|b - Ax_0\| > \|b - Ax_1\| > \|b - Ax_2\| > \ldots > \|b - Ax_i\| > \|b - Ax_{i+1}\| > \ldots$$

because the error (the 'difference' between the solution and the estimation) is going down to zero.

In practice, the algorithm stops when the initial error is divided by 100, 1000 or 10,000, depending on the overall problem to solve.

In general there is no need for x_0 to be an accurate estimation of the solution. If it is, then it usually increases the speed of the process, but if the user does not know what is required it should be set simply to the zero vector.

GMRES is in fact a direct method: by performing a defined number of operations the user will reach the exact solution. However its algorithm looks like an iterative method, and to avoid using too much memory and reach a very good performance, it is used as an iterative method. For this reason it is called a *pseudo-iterative* method. In practice, GMRES has some practical advantages over other methods: it can solve non-symmetric problems and there is no need to know precisely the A matrix. Only the result of the product of the matrix by a vector is required. For non-linear problems, where the matrix is not known but only approximated, it can be very useful.

Non-mathematicians and readers who are not interested in the derivation of the algorithm can go straight to the section on the GMRES algorithm on page 185. For others, we now describe briefly the methodology.

GMRES is a minimum residual method, the objective of which is to obtain the vector x that minimises the l^2 norm of the residual. So we seek the value x that fulfils

$$\min_{x \epsilon \Re^N} \|b - Ax\| \tag{9.13}$$

Assume $x = x_0 + z$ where z is a first estimation of x. We then look for the value z that belongs to K_k, the Krylov vectorial space. This the vectorial space whose base is $(r_0, Ar_0, \ldots, A^{k-1}r_0)$ where $r_0 = b - Ax_0$. The minimisation problem can then be written as the value of z that fulfils

$$\min_{x \epsilon K_t} \|r_0 - Az\| \tag{9.14}$$

Using a Gram–Schmidt orthogonalisation process, we build an orthonormal base $U_k = (u_1, u_2, \ldots u_k)$of K_k as well as a rectangular Hessenberg matrix H whose dimensions are $(k + 1) \times k$. H is defined by

$$AU_k = U_{k+1}H_k \tag{9.15}$$

Then $h_{i,j}$, the element of H that is on row i and column j, can be written

$$h_{i,j} = AU_j \bullet u_i \tag{9.16}$$

where • is the scalar product of two vectors.

By definition of U_k we have $h_{i,j} = 0$ when $i > j + 1$. This explains is why H is called a Hessenberg matrix. We extend their definition to rectangular matrices that have an upper triangular matrix in which we have an extra non zero under-diagonal.

In addition we have:

$$h_{j+1j} = \| u_{j+1} \| \tag{9.17}$$

z belongs to K_k and can therefore be written:

$$z = \sum_{j=1}^{k} y_j u_j \tag{9.18}$$

By defining $e \in \Re^{k+1}$ $e = (\| r_0 \|,0,...,0)^t$, it can be shown that the original minimisation problem can now be written

$$\min_{z \in K_k} \| r_0 - AZ \| \Leftrightarrow \min_{y \in \Re^k} \| e - H_k y \| \tag{9.19}$$

The quasi-triangular structure of H_k eases the process to solve such problems; we use an algorithm based on a stable QR factorisation. Above all, it gives the value of the minimal residual as a sub-product (so that there is no need for extra computation) at each iteration.

Therefore we build a matrix Q such that

$$QH_k = R \tag{9.20}$$

where R is a $(k + 1) \times k$ 'upper triangular' matrix with its last row only made of zeros, as shown below.

$$R = \begin{bmatrix} x & x & x & x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x & x & x & x \\ 0 & 0 & x & x & x & x & x & x & x & x \\ 0 & 0 & 0 & x & x & x & x & x & x & x \\ 0 & 0 & 0 & 0 & x & x & x & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x & x & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Q will be a unitary matrix, a product of k rotation matrices.

The problem can now be written as

$$\min_{y \in \Re^k} \| e - H_k y \| = \min_{y \in \Re^k} \| Qe - Ry \| \tag{9.21}$$

The solution of the problem is then obtained by solving the upper triangular system:

$$\tilde{R} y = \tilde{Q} e \tag{9.22}$$

where $\tilde{R}$ is the $(k \times k)$ matrix built using the k first rows of R, and where $\tilde{Q}e$ is the vector made of the k first components of Qe.

As we observe in the following, the error made is then given by e_{k+1}:

$$Qe - Ry = \begin{pmatrix} x \\ x \\ x \\ x \\ x \\ x \\ x \\ x \\ x \\ x \\ x \\ e_{k+1} \end{pmatrix} \begin{pmatrix} x & x & x & x & x & x & x & x & x & x \\ 0 & x & x & x & x & x & x & x & x & x \\ 0 & \mathbf{0} & x & x & x & x & x & x & x & x \\ 0 & 0 & 0 & x & x & x & x & x & x & x \\ 0 & 0 & 0 & 0 & x & x & x & x & x & x \\ 0 & 0 & 0 & 0 & 0 & x & x & x & x & x \\ 0 & 0 & 0 & \mathbf{0} & 0 & 0 & x & x & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x & x \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & x \\ 0 & 0 & 0 & 0 & 0 & \mathbf{0} & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ x \\ x \\ x \\ x \\ x \\ x \\ x \\ x \\ x \end{pmatrix}$$

GMRES algorithm

We now describe in detail the GMRES algorithm, which will allow the reader to implement it reasonably quickly.

It has already been said that the only requirement for GMRES regarding the matrix A is to be able to compute the product of A with a vector. Therefore, it is advisable to implement a standalone subroutine for GMRES that calls another routine to perform this multiplication. The algorithm will then be easily reusable to solve numerous different linear systems.

Data

A: the matrix of the linear system to solve
b: the second member of the system (vector)
k_0: the dimension of the Krylov space (set to 5 or 10)
ε: the convergence parameter (set to 10^{-4} or 10^{-5}).

Internal variable

TEST: Boolean

Output

x: solution of $Ax = b$

Begin

{Initialisation of x and of the criteria to stop}

x ← 0

TEST ← False

While *TEST* = False

Do

{Initialisation of the GMRES loop}

$u_1 \leftarrow b - Ax$

$e \leftarrow (\| u_1 \|, 0, 0, 0, \ldots 0)^t$

$$u_1 \leftarrow \frac{u_1}{\| u_1 \|}$$

{Beginning of the GMRES loop}

For i =1 to k_0

Do

{Building the Krylov vector number i+1}

$u_{i+1} \leftarrow Au_i$

For j = 1 To i

Do

$\beta_{i+1,j} \leftarrow u_{i+1} \bullet u_j$

$u_{i+1} \leftarrow u_{i+1} - \beta_{i+1,j} u_j$

End For

$$u_{i+1} \leftarrow \frac{u_{i+1}}{\| u_{i+1} \|}$$

{Building column number i of the Hessenberg Matrix H and update its QR factorisation}

$h_i \leftarrow (\beta_{i+1,1}, \ \beta_{i+1,2}, \ \beta_{i+1,i}, \| u_{i+1} \|)'$

For j = 1 To i−1

Do

$$\begin{pmatrix} h_{j,i} \\ h_{j+1,i} \end{pmatrix} \leftarrow \begin{pmatrix} \cos\theta_j & \sin\theta_j \\ -\sin\theta_j & \cos\theta_j \end{pmatrix} \begin{pmatrix} h_{j,i} \\ h_{j+1,i} \end{pmatrix}$$

End For

$$(\cos\theta_i \ \ \sin\theta_i) \leftarrow \frac{1}{\sqrt{h^2_{i,i} + h^2_{i+1,j}}} \left((h_{i,i} \ \ h_{i+1,i} \right)$$

$$(h_{i,i} \ \ h_{i+1,i}) \leftarrow (\sqrt{h^2_{i,i} + h^2_{i+1,i}} \ \ 0)$$

{update e, the error vector}

$(e_i \ \ e_{i+1}) \leftarrow (\cos\theta_i e_i \ \ \sin 0_i e_i)$

TEST ← $(|e_{i+1}| \leq \varepsilon)$

If *TEST*

Then

GOTO (**)

End If

End For
(**){compute ×}
Solve the upper triangular linear system: $Hy = e$

$$x \leftarrow x + \sum_{j=1}^{i} y_j u_j$$

End While
End

B-Splines

A key application of basis splines or *B-splines* in financial economics is when there is a need for interpolation of data. For instance when a bootstrapping methodology is used to compute zero rates from bond prices, the result is a set of rates for given dates. Obviously, if zero rates are required for other dates, an interpolation is then required to obtain these. The most popular methodology, as well as the simplest, is linear interpolation.

Example 9.4

Consider the following scenario:

- one-year zero rate of 10%
- two-year zero rate of 15%.

We require the 16-month rate Z16. Linear interpolation will give

Z16 = 10% + (15%–10%) *4/12 = 11.66%

However using linear interpolation assumes that the zero-rate yield curve is linear by time bucket. This is not an accurate representation of reality as it does not take into account the smoothness and curvature of yield curves in the real world. Therefore if we require more realistic curves, for example for use when setting up arbitrage trading strategies, a B-spline can be used.

As observed earlier, B-splines are also used for defining discount curves as a linear combination of B-splines. The advantage of using them in this context is that it is straightforward to impose constraints on the curve itself. As we saw, the discount factor for the date of calculation (today) should be 1, and the discount factor at infinite should be zero (the discount curve should go down to zero when time goes by).

Definitions

We define the following:

- Let $(t_i)_{i=0\ldots m}$ be a suite of $(m+1)$ points such as $\forall i, t_i \leq t_{i+1}$. These points are called *nodes*. If r number of t_i are equal to τ, then τ is said to be an r-order node or node of multiplicity r.

- For each (i,j) such as $1 \leq j \leq m + 1 - i$, we define $\omega_{i,j}(x)$ by:

$$\omega_{i,j}(x) = \begin{cases} \dfrac{x - t_i}{t_{i+j} - t_i} & \text{if } t_l < t_{i+j} \\ 0 \text{ in all other cases} \end{cases} \tag{9.23}$$

We can then define B-splines $B_{i,k}(x)$, $x \in \Re$ recursively on k as

$$\begin{cases} \forall i,\ 0 \leq i, \leq m-k-1 \\ B_{i,0}(x) = \begin{cases} 1 \text{ if } t_i \leq x \leq t_{i+1} \\ 0 \text{ in all other cases} \end{cases} \\ B_{i,k}(x) = \omega_{i,k}(x) B_{i,k-1}(x) + (1 - \omega_{i,+1,k}(x)) B_{i,+1,k-1}(x), \forall k \geq 1 \end{cases} \tag{9.24}$$

It can be shown that $B_{i,k}(x)$ can be written as

$$B_{i,k}(x) = \frac{x - t_i}{t_{i+k} - t_i} B_{i,k-1}(x) + \frac{t_{i+k+1} - x}{t_{i+k+1} - t_{i+1}} B_{i+1,k-1}(x), \text{ if } t_i < t_{i+k} \text{ and } t_{i+1} < t_{i+k+1} \tag{9.25}$$

By definition of $\omega_{i,k}(x)$ those terms where the denominator would be nil are set to zero.

Properties

Assume k is given:

1. $B_{i,k}(x)$ is polynomial of degree k by bucket.
2. $B_{i,k}(x) = 0$ when $x \notin [t_i, t_{i+k+1}]$
3. $B_{i,k}(x) > 0$ when $x \in [t_i, t_{i+k+1}]$
4. $B_{i,k}(t_i) = 0$ except if $t_i = t_{i+1} = \ldots = t_{i+k} < t_{i+k+1}$. Then $B_{i,k}(t_i) = 1$.
5. Given [a,b] an interval such as $t_k \leq a$ and $t_{m-k} \geq b$, then $\forall x \in [a,b]$, we have:

$$\sum_{i=0}^{m-k-1} B_{i,k}(x) = 1$$

6. Assume $x \in [t_i, t_{i+k+1}]$, then $B_{i,k}(x) = 1$ if and only if $x = t_{i+1} = \ldots = t_{i+k}$.
7. $B_{i,k}(x)$ is infinitely right-differentiable for each $x \in \Re$.

First derivative

During the optimisation process to compute a discount curve (see above), the first derivative of B-splines is required. This is:

$$B_{i,k}(x) = k \left[\frac{B_{i,k}(x)}{t_{i+k} - t_i} - \frac{B_{i+1,k-1}(x)}{t_{i+k+1} - t_{i+1}} \right] \tag{9.26}$$

We retain the convention that if a denominator is equal to zero, this means that the term itself is equal to zero.

Generating curves using B-splines

As the basis has been described, we can move on to the main issue of curve generation using B-splines. In the following we deal with cubic splines (when $k = 3$). Although any number may be used in a polynomial spline application, generally cubic splines are the ones most commonly used in financial applications.

Definitions

- We set n points $(P_i)_{i=0\ldots n-1}$ of $\Re^3$. γ is the parameterised B-spline curve associated to the polygon $P_0 \ldots P_{n-1}$, defined by

$$S(t) = \begin{pmatrix} S_1(t) \\ S_2(t) \\ S_3(t) \end{pmatrix} = \sum_{i=0}^{n-1} P_i B_{i,3}(t) \ (a \le t \le b)$$

- The polygon $P_0 \ldots P_{n-1}$ is called the control polygon of the γ curve.

Properties

Properties from the B-spline functions trigger the following properties for B-spline curves:

1. In general, γ does not go through the points P_i, but if $a = t_0 = \ldots = t_3$, and $t_n = \ldots = t_{n+3} = b$, then $S(a) = P_0$ and $S(b) = P_{n-1}$. In this case, γ is tangent in P_0 and P_{n-1} to the edges (P_0, P_1) and (P_{n-2}, P_{n-1}) of the control polygon.
2. γ is in the convex envelope of the points $P_0, \ldots P_{n-1}$. More precisely, if $t_i \le t \le t_{i+1}$ then $S(t)$ is in the convex envelop of the points $P_{i-k}, \ldots P_1$.
3. If for all i such $4 \le i \le n - 1$ as the nodes t_i are simple (they are one-order nodes), then γ is of class C^2 and is made of n parameterised polynomial arcs of degree equal to or less than 3.

This generates the curve.

Conclusion

We have described an advanced methodology by which one can extract a zero-coupon curve from market-observed bond prices. It is more reliable than the classic bootstrapping methodology as it smoothes the discount curve, and therefore the zero-curve, thanks to its B-spline definition. More importantly, it results in more realistic forward rates.

We also described the basic tools, that is, B-splines and the optimisation method, so that non-mathematicians should be able to implement this methodology without undue complication and perhaps on their own.

Selected bibliography and references

Brown, P. N. and Saad, Y. 'Hybrid Krylov method for non linear systems of equations', *Journal of Statistical Computation*, 11, 3, May 1990.

De Clermont-Tonnerre, A., and Lévy, M. A. *Zero-Coupon Bonds and Bond Stripping*, IFR Publishing, 1995.

James, N. and Webber, A. *Interest Rate Modelling*, Wiley, 2000.

Joannas, D. *Optimisation de formes en aérodynamique,* Ph.D. thesis, Université de Saint Etienne, France, 1992.

Vasicek, O., Oldrich, A. and Gifford-Fong, H. Term structure modelling using exponential splines, *Journal of Finance*, 37, 2, May 1982.

Riesler, J. J. *Méthodes mathématiques pour la C.A.O.*, Masson, 1991.

10
Inflation-Indexed Bonds and Derivatives

In certain countries there is a market in bonds whose return, both coupon and final redemption payment, is linked to the consumer prices index. Investors' experience with inflation-indexed bonds differs across countries, as they were introduced at different times, and as a result the exact design of *index-linked bonds* varies across the different markets. This of course makes the comparison of issues such as yield difficult, and has in the past acted as a hindrance to arbitrageurs seeking to exploit real yield differentials. In this chapter we highlight the basic concepts behind the structure of indexed bonds and show how these differ from those employed in other markets. Not all index-linked bonds link both coupon and maturity payments to a specified index; in some markets only the coupon payment is index-linked. Generally the most liquid market available will be the government bond market in index-linked instruments.

The structure of index-linked bond markets across the world varies in various ways including those noted below. Appendix 10.1 lists those countries that currently issue public-sector indexed securities.

Introduction and basic concepts

There are a number of reasons that investors and issuers alike are interested in inflation-indexed bonds. Before considering these, we look at some of the factors involved in security design.

Choice of index

In principle bonds can be indexed to any number of variables, including various price indices, earnings, output, specific commodities or foreign currencies. Although ideally the chosen index would reflect the hedging requirements of both parties, these may not coincide. For instance the overwhelming choice of retail investors is for indexation to consumer prices, whereas pension funds prefer linking to earnings levels, to offset earnings-linked pension liabilities. In practice most bonds have been linked to an index of consumer prices such as the UK Retail Price Index, since this is usually widely circulated and well understood and issued on a regular basis.

Indexation lags

In order to provide practically precise protection against inflation, interest payments for a given period would need to be corrected for actual inflation over the same period. However for two important reasons, there are unavoidable lags between the movements in the price index and the adjustment to the bond cash flows as they are paid. This reduces the inflation-proofing properties of indexed bonds. Deacon and Derry (1998) state two reasons that indexation lags are necessary. First, inflation statistics can only be calculated and published with a delay. The data for one month is usually only known well into the next month, and there may be delays in publication. This calls for a lag of at least one month. Secondly, in some markets the size of the next coupon payment must be known before the start of the coupon period in order to calculate the accrued interest; this leads to a delay equal to the length of time between coupon payments.[1]

The indexation lag is illustrated at Figure 10.1.

Coupon frequency

Index-linked bonds often pay interest on a semi-annual basis, but long-dated investors such as fund managers whose liabilities may well include inflation-indexed annuities are also, at least in theory, interested in indexed bonds that pay on a quarterly or even monthly basis.

Indexing the cash flows

There are four basic methods of linking the cash flows from a bond to an inflation index. These are:

- *Interest-indexed bonds*. These pay a fixed real coupon and an indexation of the fixed principal every period; the principal repayment at maturity is not adjusted. In this case all the inflation adjustment is fully paid out as it occurs

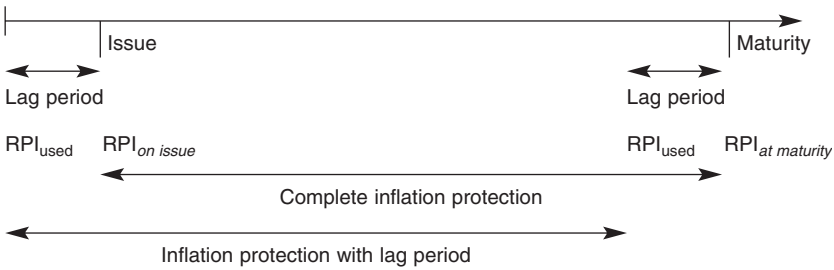


Figure 10.1 The indexation lag

[1]The same source cites various methods by which the period of the lag may be minimised: for example the accrued interest calculation for Canadian Real Return Bonds is based on cumulative movements in the consumer prices index, which run from the last coupon date. This obviates the need to know with certainty the nominal value of the next coupon, unlike the arrangement for UK index-linked gilts. (See Deacon and Derry, 1998, pp. 30–1).

and does not accrue on the principal. This type of bonds has been issued in Australia, although the most recent issue was in 1987.

- *Capital-indexed bonds*. The coupon rate is specified in real terms. Interest payments equal the coupon rate multiplied by the inflation-adjusted principal amount. At maturity the principal repayment is the product of the nominal value of the bond multiplied by the cumulative change in the index. Compared with interest-indexed bonds of similar maturity, these bonds have higher duration and lower reinvestment risk. This type of bonds has been issued in Australia, Canada, New Zealand, the UK and the United States.
- *Zero-coupon indexed bonds*. As their name implies these pay no coupons but the principal repayment is scaled for inflation. These have the highest duration of all indexed bonds and have no reinvestment risk. This type of bonds has been issued in Sweden.
- *Indexed-annuity bonds*. The payments consist of a fixed annuity payment and a varying element to compensate for inflation. These bonds have the lowest duration and highest reinvestment risk of all index-linked bonds. They have been issued in Australia, although not by the central government.
- *Current pay bond*. As with interest indexed bonds, the principal cash flow on maturity is not adjusted for inflation. The difference with current pay bonds is that their term cash flows are a combination of an inflation-adjusted coupon and an indexed amount that is related to the principal. Thus in effect current pay bonds are an inflation-indexed floating-rate note. They have been issued in Turkey.

The choice of instrument will reflect the requirements of investors and issuers. Deacon and Derry (1998) cite duration, tax treatment and reinvestment risk as the principal factors that influence instrument design. Although duration for an indexed bond measures something slightly different from that for a conventional bond, being an indication of the bond price sensitivity due to changes in the real interest rate, as with conventional bonds it is higher for zero-coupon indexed bonds than for coupon bonds. Indexed annuities will have the shortest duration. Longer-duration instruments will (in theory) be demanded by investors that have long-dated hedging liabilities. Again similarly to conventional bonds, investors holding indexed bonds are exposed to reinvestment risk, which means that the true yield earned by holding a bond to maturity cannot be determined when it is purchased, as the rate at which interim cash flows can be invested is not known. Hence bonds that pay more of their return in the form of coupons (such as indexed annuities) are more exposed to this risk. Indexed zero-coupon bonds, like their conventional counterparts, do not expose investors to reinvestment risk. The tax regime in individual markets will also influence investor taste. For instance some jurisdictions tax the capital gain on zero-coupon bonds as income, with a requirement that any tax liability be discharged as current income. This is unfavourable treatment as the capital is not available until maturity, which would reduce institutional investor demand for zero-coupon instruments.

It should be noted also that in three countries, namely Canada, New Zealand and the United States, there exists a facility for investors to strip indexed bonds,

thus enabling separate trading of coupon and principal cash flows.[2] Such an arrangement obviates the need for a specific issue of zero-coupon indexed securities, as the market can create them in response to investor demand.

Coupon stripping feature

Allowing market practitioners to strip indexed bonds enables them to create new inflation-linked products that are more specific to investors needs, such as indexed annuities or deferred payment indexed bonds. In markets that allow stripping of indexed government bonds, a strip is simply an individual uplifted cash flow. An exception to this is in New Zealand, where the cash flows are separated into three components, the principal, the principal inflation adjustment and the set of inflation-linked coupons (that is, an indexed annuity).

Index-linked bond yields

Calculating index-linked yields

Inflation-indexed bonds have either or both of their coupon and principal linked to a price index such as a retail price index (RPI), a commodity price index (for example, wheat) or a stock market index. In the UK the reference is to the RPI whereas in other markets the price index is the consumer price index (CPI). If we wish to calculate the yield on such bonds it is necessary to make forecasts of the relevant index, which are then used in the yield calculation. In the UK both the principal and coupons on UK index-linked government bonds are linked to the RPI and are therefore designed to give a constant real yield. Most of the index-linked stocks that have been issued by the UK government have coupons of 2% or 2.5%. This is because the return from an index-linked bond in theory represents real return, as the bond's cash flows rise in line with inflation. Historically the real rate of return on UK market debt stock over the long-term has been roughly 2.5%.

Indexed bonds differ in their make-up across markets. In some markets only the principal payment is linked, whereas other indexed bonds link only their coupon payments and not the redemption payment. In the case of the former, each coupon and the final principal are scaled up by the ratio of two values of the RPI. The main RPI measure is the one reported for eight months before the issue of the gilt, and is known as the *base RPI*. The base RPI is the denominator of the index measure. The numerator is the RPI measure for eight months prior to the month coupon payment, or eight months before the bond maturity date.

The coupon payment of an index-linked gilt is given by (10.1):

$$\text{Coupon payment} = (\text{C}/2) \times \frac{RPI_{C-8}}{RPI_0} \tag{10.1}$$

[2]In the UK the facility of 'stripping' exists for conventional gilts but not index-linked gilts. The term originates in the US market, being an acronym for Separate Trading of Registered Interest and Principal.

This expression shows the coupon divided by two before being scaled up, because index-linked gilts pay semi-annual coupons. The formula for calculating the size of the coupon payment for an annual-paying indexed bond is modified accordingly.

The principal repayment is given by (10.2):

$$\text{Principal repayment} = 100 \times \frac{RPI_{M-8}}{RPI_0} \tag{10.2}$$

where

C is the annual coupon payment
RPI_0 is the RPI value eight months prior to the issue of the bond (the *base* RPI)
RPI_{C-8} is the RPI value eight months prior to the month in which the coupon is paid
RPI_{M-8} is the RPI value eight months prior to the bond redemption.

Price indices are occasionally 're-based', which means that the index is set to a base level again. In the UK the RPI has been re-based twice, the last occasion being in January 1987, when it was set to 100 from the January 1974 value of 394.5.

Example 10.1

An index-linked gilt with coupon of 4.625% was issued in April 1988 and matured in April 1998. The base measure required for this bond is the RPI for August 1987, which was 102.1. The RPI for August 1997 was 158.5. We can use these values to calculate the actual cash amount of the final coupon payment and principal repayment in April 1998, as shown below.

$$\text{Coupon payment} = (4.625/2) \times \frac{158.5}{102.1} = £3.58992$$

$$\text{Principal repayment} = 100 \times \frac{158.5}{102.1} = £14.23996$$

We can determine the accrued interest calculation for the last six-month coupon period (October 1987 to April 1998) by using the final coupon payment, given below.

$$3.58992 \times \frac{\text{No. of days accured}}{\text{actual days in period}}$$

The markets use two main yield measures for index-linked bonds, both of which are a form of yield-to-maturity. These are the *money* (or *nominal*) *yield*, and the *real yield*.

In order to calculate a money yield for an indexed bond we require forecasts of all future cash flows from the bond. Since future cash flows from an index-linked bond are not known with certainty, we require a forecast of all the relevant future RPIs, which we then apply to all the cash flows. In fact the market convention is to take the latest available RPI and assume a constant inflation rate thereafter, usually 2.5% or 5%. By assuming a constant inflation rate we can set future RPI levels, which in turn allow us to calculate future cash flow values.

We obtain the forecast for the first relevant future RPI using (10.3).

$$RPI_1 = RPI_0 \times (1+\tau)^{m/12} \tag{10.3}$$

where

RPI_1 is the forecast RPI level
RPI_0 is the latest available RPI
τ is the assumed future annual inflation rate
m is the number of months between RPI_0 and RPI_1.

Consider an indexed bond that pays coupons every June and December. For analysis we require the RPI forecast value for eight months prior to June and December, which will be for October and April. If we are now in February, we require a forecast for the RPI for the next April. This sets $m = 2$ in our equation at (10.3). We can then use (10.4) to forecast each subsequent relevant RPI required to set the bond's cash flows.

$$RPI_{j+1} = RPI_1 \times (1+\tau)^{j/2} \tag{10.4}$$

where j is the number of semi-annual forecasts after RPI_1 (which was our forecast RPI for April). For example if the February RPI was 163.7 and we assume an annual inflation rate of 2.5%, then we calculate the forecast for the RPI for the following April to be:

$$RPI_1 = 163.7 \times (1.025)^{2/12} = 164.4$$

and for the following October it would be:

$$RPI_3 = 164.4 \times (1.025) = 168.5$$

Once we have determined the forecast RPIs we can calculate the yield. Under the assumption that the analysis is carried out on a coupon date, so that accrued interest on the bond is zero, we can calculate the money yield (ri) by solving equation (10.5).

$$P_d = \frac{(C/2)(RPI_1/RPI_0)}{(1+\frac{1}{2}ri)} + \frac{(C/2)(RPI_2/RPI_0)}{(1+\frac{1}{2}ri)^2} + \ldots + \frac{([C/2]+M)(RPI_N/RPI_0)}{(1+\frac{1}{2}ri)^N} \tag{10.5}$$

where

ri is the semi-annualised money yield-to-maturity
N is the number of coupon payments (interest periods) up to maturity.

Equation (10.5) is for semi-annual paying indexed bonds such as index-linked gilts. The equation for annual coupon indexed bonds is given at (10.6).

$$P_d = \frac{C(RPI_1/RPI_0)}{(1+ri)} + \frac{C(RPI_2/RPI_0)}{(1+ri)^2} + ... + \frac{(C+M)(RPI_N/RPI_0)}{(1+ri)^N} \tag{10.6}$$

The real yield *ry* is related to the money yield through equation (10.7), as it applies to semi-annual coupon bonds. It was first described by Fisher in *Theory of Interest* (1930).

$$(1+\tfrac{1}{2}ry) = (1+\tfrac{1}{2}ri)/(1+\tau)^{\frac{1}{2}} \tag{10.7}$$

To illustrate this, if the money yield is 5.5% and the forecast inflation rate is 2.5%, then the real yield is calculated using (10.7) as shown below.

$$ry = \left\{\frac{[1+\frac{1}{2}(0.055)]}{[1+(0.025)]^{\frac{1}{2}}} - 1\right\} \times 2 = 0.0297 \text{ or } 2.97\%$$

We can rearrange equation (10.5) and use (10.7) to solve for the real yield, shown at (10.8) and applicable to semi-annual coupon bonds. Again, we use (10.8) where the calculation is performed on a coupon date.

$$\begin{aligned} P_d &= \frac{RPI_a}{RPI_0}\left[\frac{(C/2)(I=\tau)^{\frac{1}{2}}}{(1+\frac{1}{2}ri)} + \frac{(C/2)(I+\tau)}{(1+\frac{1}{2}ri)^2} + \ldots + \frac{(\{C/2\}+M)(I+\tau)^{\frac{N}{2}}}{(1+\frac{1}{2}ri)^N}\right] \\ &= \frac{RPI_a}{RPI_0}\left[\frac{(C/2)}{(1+\frac{1}{2}ry)} + \ldots + \frac{(C/2)+M}{(1+\frac{1}{2}ry)^N}\right] \end{aligned} \tag{10.8}$$

where

$$PRI_a = \frac{PRI_1}{(1+\tau)^2}$$

RPI_0 is the base index level as initially described. RPI_a/RPI_0 is the rate of inflation between the bond's issue date and the date the yield calculation is carried out.

It is best to differentiate between the equations for money yield and real yield by thinking of which discount rate to employ when calculating a redemption yield for an indexed bond. Equation (10.5) can be viewed as showing that the

money yield is the appropriate discount rate for discounting money or nominal cash flows. We then re-arrange this equation as given in (10.8) to show that the real yield is the appropriate discount rate to use when discounting real cash flows.

The yield calculation for US Treasury inflation-indexed securities is given at Appendix 10.2.

Assessing yield for index-linked bonds

Index-linked bonds do not offer *complete* protection against a fall in the real value of an investment. That is, the return from index-linked bonds including index-linked gilts is not in reality a guaranteed real return, in spite of the cash flows being linked to a price index such as the RPI. The reason for this is the lag in indexation, which for index-linked gilts is eight months. The time lag means that an indexed bond is not protected against inflation for the last interest period of its life, which for gilts is the last six months. Any inflation occurring during this final interest period will not be reflected in the bond's cash flows and will reduce the real value of the redemption payment and hence the real yield. This can be a worry for investors in high-inflation environments. The only way to effectively eliminate inflation risk for bondholders is to reduce the time lag in indexation of payments to something like one or two months.

Bond analysts frequently compare the yields on index-linked bonds with those on conventional bonds, as these reflect the market's expectation of inflation rates. To compare returns between index-linked bonds and conventional bonds analysts calculate the *break-even inflation rate*. This is the inflation rate that makes the money yield on an index-linked bond equal to the redemption yield on a conventional bond of the same maturity. Roughly speaking the difference between the yield on an indexed bond and a conventional bond of the same maturity is what the market expects inflation to be during the life of the bond; part of the higher yield available on the conventional bond is therefore the inflation *premium*. In August 1999 the redemption yield on the 5¾% Treasury 2009, the ten-year benchmark gilt, was 5.17%. The real yield on the 2½% index-linked 2009 gilt, assuming a constant inflation rate of 3%, was 2.23%. Using (10.5) this gives us an implied break-even inflation rate of:

$$t = \left\{\frac{[1 + \frac{1}{2}(0.0517)]}{[1 + \frac{1}{2}(0.0223)]}\right\}^2 - 1 = 0.029287 \text{ or } 2.9\%$$

If we accept that an advanced, highly developed and liquid market such as the gilt market is of at least semi-strong form, if not strong form, then the inflation expectation in the market is built into these gilt yields. However if this implied inflation rate understated what was expected by certain market participants, investors would start holding more of the index-linked bond rather than the conventional bond. This activity would then force the indexed yield down (or the conventional yield up). If investors had the opposite view and thought that the implied inflation

rate overstated inflation expectations, they would hold the conventional bond. In our illustration above, the market is expecting long-term inflation to be at around 2.9% or less, and the higher yield of the 5¾% 2009 bond reflects this inflation expectation. A fund manager will take into account her view of inflation, amongst other factors, in deciding how much of the index-linked gilt to hold compared with the conventional gilt. It is often the case that investment managers hold indexed bonds in a portfolio against specific index-linked liabilities, such as pension contracts that increase their payouts in line with inflation each year.

The premium on the yield of the conventional bond over that of the index-linked bond is therefore compensation against inflation to investors holding it. Bondholders will choose to hold index-linked bonds instead of conventional bonds if they are worried by unexpected inflation. An individual's view on expected inflation will depend on several factors, including the current macro-economic environment and the credibility of the monetary authorities, whether it is the central bank or the government. In certain countries such as the UK and New Zealand, the central bank has explicit inflation targets and investors may feel that over the long term these targets will be met. If the track record of the monetary authorities is proven, investors may feel further that inflation is no longer a significant issue. In these situations the case for holding index-linked bonds is weakened.

The real-yield level on indexed bonds in other markets is also a factor. As capital markets around the world have become closely integrated in the last 20 years, global capital mobility means that high-inflation markets are shunned by investors. Therefore over time expected returns, certainly in developed and liquid markets, should be roughly equal, so that real yields are at similar levels around the world. If we accept this premise, we would then expect the real yield on index-linked bonds to be at approximately similar levels, whatever market they are traded in. For example we would expect indexed bonds in the UK to be at a level near to that in, say, the US market. In fact in May 1999 long-dated index-linked gilts traded at just over 2% real yield, while long-dated indexed bonds in the United States were at the higher real yield level of 3.8%. This was viewed by analysts as reflecting that international capital was not as mobile as had previously been thought, and that productivity gains and technological progress in the US economy had boosted demand for capital there to such an extent that real yield had had to rise. However there is no doubt that there is considerable information content in index-linked bonds, and analysts are always interested in the yield levels of these bonds compared with conventional bonds.

Analysis of real interest rates[3]

Observation of trading patterns in a liquid market in inflation-indexed bonds enables analysts to draw conclusions on nominal versus real interest indicators, and the concept of an inflation term structure. However such analysis is often problematic as there is usually a significant difference between liquidity levels of conventional and

[3]This section follows the approach adopted in Deacon and Derry (1998), chapter 5.

indexed bonds. Nevertheless, as we discussed in the previous section, it is usually possible to infer market estimates of inflation expectations from the yields observed on indexed bonds, when compared with conventional yields.

Inflation expectations

Where an indexed bond incorporates an indexation lag there is an imperfect indexation, and the bond's return will not be completely inflation-proof. Deacon and Derry (1998) suggest that this means an indexed bond may be regarded as a combination of a true indexed instrument (with no lag) and an unindexed bond. Where the lag period is exactly one coupon period the price/yield relationship is given by

$$P = \sum_{j=1}^{n} \frac{C\prod_{i=0}^{j-1}(1+ri_i)}{(1+rm_j)^j \prod_{i=1}^{j}(1+ri_i)} + \frac{M\prod_{i=0}^{n-1}(1+ri_i)}{(1+rm_n)^n \prod_{i=1}^{n}(1+ri_i)} \tag{10.9}$$

where

ri is the rate of inflation between dates *i*–1 and *i*
rm is the redemption yield

and *C* and *M* are coupon and redemption payments as usual. If the bond has just paid the last coupon ahead of its redemption date, (10.9) reduces to

$$P = \frac{C}{(1+rm)(1+ri)} + \frac{M}{(1+rm)(1+ri)} \tag{10.10}$$

In this situation the final cash flows are not indexed and the price/yield relationship is identical to that of a conventional bond. This fact enables us to quantify the indexation element, as the yields observed on conventional bonds can be compared with those on the non-indexed element of the indexed bond. This implies a true real yield measure for the indexed bond.

The Fisher identity is used to derive this estimate. Essentially this describes the relationship between nominal and real interest rates, and in one form is given as:

$$1 + y = (1+r)(1+i)(1+\rho) \tag{10.11}$$

where y is the nominal interest rate, r the real interest rate, i the expected rate of inflation, and ρ is a premium for the risk of future inflation. Using (10.11), assuming a value for the risk premium ρ can link the two bond price equations, which can, as a set of simultaneous equations, be used to obtain values for the real interest rate and the expected inflation rate.

One approach is to use two bonds of identical maturity, one conventional and one indexed, if they exist, and ignoring lag effects use the yields on both to determine the expected inflation rate, given by the difference between the redemption yields of each bond. In fact as we noted in the previous section this measures the average

expected rate of inflation during the period from now to the maturity of the bonds. This is at best an approximation. It is a flawed measure because an assumption of the expected inflation rate has been made when calculating the redemption yield of the indexed bond in the first place. As Deacon and Derry (1998, p. 91) state, this problem is exacerbated if the maturity of both bonds is relatively short, because impact of the unindexed element of the indexed bond is greater with a shorter maturity. To overcome this flaw, a *breakeven* rate of inflation is used. This is calculated by first calculating the yield on the conventional bond, followed by the yield on the indexed bond using an assumed initial inflation rate. The risk premium ρ is set to an assumed figure, say 0. The Fisher identity is used to calculate a new estimate of the expected inflation rate i. This new estimate is used to recalculate the yield on the indexed bond, which is then used to produce a new estimate of the expected inflation rate. The process is repeated iteratively until a consistent value for i is obtained.

The main drawback with this basic technique is that it is rare for there to exist a conventional and an index-linked bond of identical maturity, so approximately similar maturities have to be used, further diluting the results. The yields on each bond will also be subject to liquidity, taxation, indexation and other influences. There is also no equivalent benchmark (or *on-the-run*) indexed security. The bibliography cites some recent research that has investigated this approach.

An inflation term structure

Where a liquid market in indexed bonds exists, across a reasonable maturity term structure, it is possible to construct a term structure of inflation rates. In essence this involves fitting the nominal and real interest-rate term structures, the two of which can then be used to infer an inflation term structure. This in turn can be used to calculate a forward expected inflation rate for any future term, or a forward inflation curve, in the same way that a forward interest rate curve is constructed.

The Bank of England uses an iterative technique to construct a term structure of inflation rates.[4] First the nominal interest rate term structure is fitted using a version of the Waggoner model (1997, also described in James and Webber, 2000). An initial assumed inflation term structure is then used to infer a term structure of real interest rates. This assumed inflation curve is usually set flat at 3% or 5%. The real interest rate curve is then used to calculate an implied real interest rate forward curve. Second, the Fisher identity is applied at each point along the nominal and real interest rate forward curves, which produces a new estimate of the inflation term structure. A new real interest rate curve is calculated from this curve. The process is repeated until a single consistent inflation term structure is produced.

Inflation-indexed derivatives

Inflation-indexed derivatives, also known as inflation-linked derivatives or *inflation derivatives*, have become widely traded instruments in the capital markets in a

[4]This is a term structure of expected inflation rates.

relatively short space of time. They are traded generally by the same desks in investment banks that trade inflation-linked sovereign bonds, which use these instruments for hedging as well as to meet the requirements of clients such as hedge funds, pension funds and corporates. They are a natural development of the inflation-linked bond market.

Inflation derivatives are an additional means by which market participants can have an exposure to inflation-linked cash flows. They can also improve market liquidity in inflation-linked products, as an earlier generation of derivatives did for interest rates and credit risk. As flexible OTC products, inflation derivatives offer advantages over cash products in certain circumstances. They provide:

- an ability to tailor cash flows to meet investors' requirements
- a means by which inflation-linked exposures can be hedged
- an instrument via which relative value positions can be put on across cash and synthetic markets
- a building block for the structuring of more complex and hybrid products.

The inflation derivatives market in the UK was introduced after the introduction of the gilt repo market in 1996. In most gilt repo trades, index-linked gilts could be used as collateral; this mean that both index-linked and conventional gilts could be used as hedging tools against positions in inflation derivatives. In the euro area, index-linked derivatives were introduced later but experienced significant growth during 2002–3. The existence of a sovereign index-linked bond market can be thought of as a necessary precursor to the development of index-linked derivatives, and although there is no reason that this should be the case, up to now it has been the case. The reason for this is probably that such a cash market suggests that investors are aware of the attraction of index-linked products, and wish to invest in them. From a market in index-linked bonds then develops a market in index-linked swaps, which are the most common index-linked derivatives. The index-linked bond market also provides a ready reference point from which index-linked derivatives can be priced.

Market instruments

We describe first some common inflation derivatives, before considering some uses for hedging and other purposes. We then consider index-linked derivatives pricing.

Inflation-linked bond swap

This is also known as a *synthetic index-linked bond*. It is a swap with the following two cash flow legs:

- pay (receive) the cash flows on a government IL bond
- receive (pay) a fixed or floating cash flow.

This converts existing conventional fixed or floating-rate investments into inflation-linked investments. An example of such a swap is given in Example 10.2.

Example 10.2: Index-linked bond swap

Nominal	€100,000,000
Start date	15 March 2004
Maturity term	15 March 2009
Bank receives	Six-month Euribor flat [+ spread], semi-annual, act/360 or Fixed rate coupon x%, annual 30/360
Bank pays	Real coupon of y% y * [$HICP(p-3)/HICP(s-3)$] * *daycount* * *notional* annual 30/360 On maturity: *Notional* * max {0%,[$HICP(m-3)/HICP(s-3)-1$]}

The symbols in the formulae above are

p	payment date
s	start date
m	maturity date
HICP	Harmonised Index of Consumer Prices.

The 'minus 3' in the formula for HICP refers to a three-month lag for indexation, common in euro sovereign index-linked bond markets.
The swap is illustrated in Figure 10.2.

Year-on-year inflation swap

This swap is commonly used to hedge issues of index-linked bonds. The swap is comprised of:

- Pay (receive) an index-linked coupon, which is a fixed rate component plus the annual rate of change in the underlying index.
- Receive (pay) Euribor or Libor, plus a spread if necessary.

With these swaps, the index-linked leg is usually set at a floor of 0% for the annual change in the underlying index. This guarantees the investor a minimum return of the fixed rate coupon.

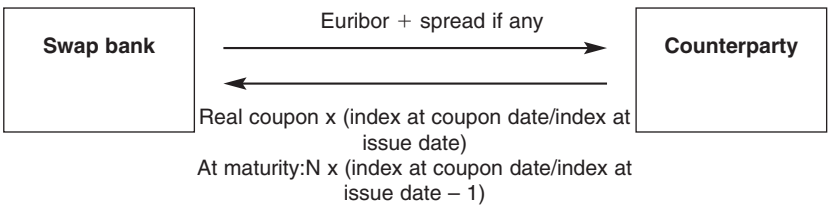


Figure 10.2 Synthetic index-linked bond

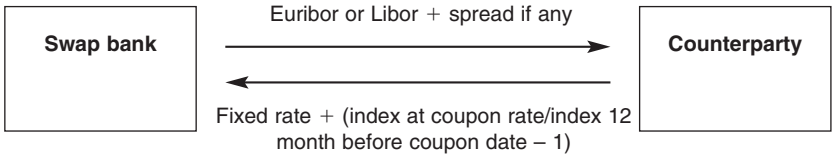


Figure 10.3 Year-on-year inflation swap

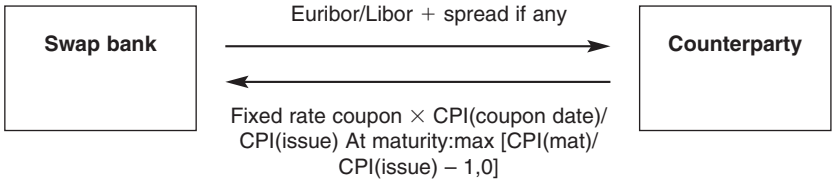


Figure 10.4 Illustration of a TIPS swap

This swap is also known as a *pay-as-you-go* swap. It is shown in Figure 10.3.

TIPS swap

The TIPS swap is based on the structure of US inflation-indexed (TIPS) securities. It pays a periodic fixed rate on an accreting notional amount, together with an additional one-off payment on maturity. This payout profile is identical to many government index-linked bonds. They are similar to synthetic index-linked bonds described above.

TIPS swaps are commonly purchased by pension funds and other long-dated investors. They may prefer the added flexibility of the index-linked swap market compared with the cash index-linked bond market. Figure 10.4 shows the TIPS swap.

Breakeven swap

This is also known as a *zero-coupon inflation swap* or *zero-coupon swap*. It allows the investor to hedge away a breakeven exposure. Compared with index-linked swaps such as the synthetic bond swap, which hedge a real yield exposure, the breakeven swap has both cash flow legs paying out on maturity. The legs are:

- the total return on the inflation index
- a compounded fixed breakeven rate.

This structure enables index-linked derivative market makers to hedge their books. It is illustrated in Figure 10.5.

Real annuity swap

A real annuity swap is used to hedge inflation-linked cash flows where this applies for payments such as rental streams, lease payments and project finance cash flows. It enables market participants who pay or receive such payments to replace

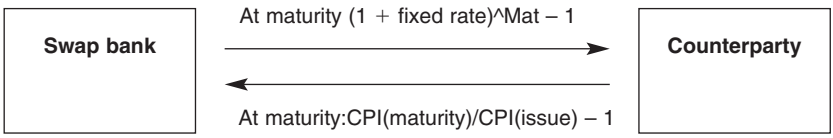


Figure 10.5 Breakeven inflation swap

the uncertainty of the future level of these cash flows with a fixed rate of growth. The swap is written on the same notional amount for both legs, but payout profiles differ as follows:

- The index-linked leg of the swap compounds its payments with the rate of change of the index.
- The fixed leg of the swap compounds its payments at a prespecified fixed rate.

These swaps are one of the most commonly traded. The fixed rate quoted for the swap provides a ready reference point against which to compare expected future rates of inflation. So for instance, if a bank is quoting for a swap with a fixed rate of 3%, and an investor believes that inflation rates will not rise above 3% for the life of the swap, the investor will receive 'fixed' (here meaning a fixed rate of growth) and pay inflation-linked on the swap.

The inflation term structure and pricing inflation derivatives

An inflation term structure is a necessary prerequisite to the pricing of inflation derivatives. It is constructed using the same principles we discussed in Chapters 3 and 4. Previously, to construct this curve we would have used index-linked bond prices as the set of market yields used as inputs to the curve. Now however we can also use the prices of index-linked derivatives. As with other markets, the derivative prices are often preferred to cash prices for two reasons: first, we can use a continuous set of prices rather than have to rely on available bond maturities, and second, there is usually greater liquidity in the over-the-counter (OTC) market.

In the case of index-linked products, the indexation element is not in fact a true picture, but rather a picture based on a lag of three, six or seven months. This lag needs to be taken into account when constructing the curve.

The forward index value I at time T from time $t(t < T)$ is given by:

$$I(t,T) = \frac{I(t)P_r(t,T)}{P_n(t,T)} \tag{10.12}$$

where

$I(t)$ is the index value at time t
$P_r(t,T)$ is the price at time t of a real zero-coupon bond of par value 1 maturing at T
$P_n(t,T)$ is the price at time t of a nominal zero-coupon bond of par value 1 maturing at T.

Using equation (10.12) we can build a forward inflation curve provided we have the values of the index at present, as well as a set of zero-coupon bond prices of required credit quality. Following standard yield curve analysis, we can build the term structure from forward rates and therefore imply the real yield curve, or alternatively we can construct the real curve and project the forward rates. However if we are using inflation swaps for the market price inputs, the former method is preferred because index-linked swaps are usually quoted in terms of a forward index value. The curve can be constructed using standard bootstrapping techniques.

Inflation derivatives can be priced reasonably accurately once the inflation term structure is constructed. However some practitioners use stochastic models in pricing such products to account for the volatility surfaces. That is, they model the volatility of inflation as well. The recent literature describes such methods. For instance, van Bezooyen, Exley and Smith (1997), Hughston (1998) and Jarrow and Yildirm (2003) suggest an approach based on that described by Amin and Jarrow (1991). This assumes that suitable proxies for the real and nominal term structures are those of foreign and the domestic economies. In other words, the foreign exchange rate captures the information required to model the two curves. We describe this approach here.

We assume that the index follows a lognormal distribution, and we use normal models for the real and nominal forward rates. For the index we then have:

$$dI(t) = I(t)[\mu_i(t)dt + \sigma(t)dW(t)] \tag{10.13}$$

and we have

$$dF_n(t,T) = \alpha_n(t,T)dt + \sigma_n(t,T)dW(t) \tag{10.14}$$

$$dF_r(t,T) = \alpha_r(t,T)dt + \sigma_r(t,T)dW(r) \tag{10.15}$$

for the nominal and real forward rate processes.

The dynamics of the zero-coupon bonds introduced earlier for equation (10.12) are given by

$$d(\log(P_k(t,T))) = \left[r_k(t) - \int_t^T a_k(t,u)du\right]dt + \sum_k(t,T)dw(t) \tag{10.16}$$

where $k = n, r$ and where

$$\sum_k(t,T) = -\int_t^T s_k(t,u)du$$

$$k = n, r$$

describes the zero-coupon bond volatilities. We use a one-factor model (see Chapter 4) for each of the term structures and one for the index. Therefore $dW(t)$

is a combined three-dimensional vector of three correlated Brownian or Weiner or processes, with a correlation of ρ. The volatility of each bond and the index is therefore also a three-element vector.

From the above, the price at time t of an option on the index struck at X and expiring at time T is given by:

$$V_{\sigma}^{Index}(t,T) = \phi[I(t)P_r(t,T)N(\phi h_1) - XP_n(t,T)N(\phi h_2)] \tag{10.17}$$

where $\phi = 1$ for a call option and $\phi = -1$ for a put option, and where

$$h_1 = \left(\log\left(\frac{I(t)P_r(t,T)}{P_n(t,T)} \right) + \frac{V(t,T)^2}{2} \right) / V(t,T)$$

$$h_2 = h_1 - V(t,T)$$

and where

$$V(t,T)^2 = \int_t^T \sum_I (u,T) \cdot \rho \cdot \sum_I (u,T) du$$

and where we define

$$\sum_I (u,T) = \sum_n (u,T) - \sum_r (u,T) - \sigma_I(u)$$

The result above has been derived, in different forms, in all three references noted above. As with other options pricing models, it needs to be calibrated to the market before it can be used. Generally this will involve using actual and project forward inflation rates to fit the model to market prices and volatilities.

Applications

We now describe some common applications of index-linked derivatives.

Hedging pension liabilities

This is perhaps the most obvious application. Assume a life assurance company or corporate pension fund wishes to hedge its long-dated pension liabilities, which are linked to the rate of inflation. It may invest in sovereign index-linked bonds such as index-linked gilts, or in index-linked corporate bonds that are hedged (for credit risk purposes) with credit derivatives. However the market in index-linked bonds is not always liquid, especially in index-linked corporate bonds. The alternative is to buy a synthetic index-linked bond. This is structured as a combination of a conventional government bond and an index-linked swap, in which the pension fund pays away the bond coupon and receives inflation-linked payments.

The net cash flow leaves the pension fund receiving a stream of cash flow that is linked to inflation. The fund is therefore hedged against its liabilities. In addition, because the swap structure can be tailor-made to the pension fund's requirements, the dates of cash flows can be set up exactly as needed. This is an added advantage over investing in the index-linked bonds directly.

Portfolio restructuring using inflation swaps

Assume that a bank or corporate has an income stream that is linked to inflation. Up to now, it has been funded by a mix of fixed and floating-rate debt. Say that these are floating-rate bank loans and fixed-rate bonds. However from an asset–-liability management (ALM) point of view this is not optimal, because of the nature of a proportion of its income. It makes sense, therefore, to switch a part of its funding into an inflation-linked segment. This can be done using either of the following approaches:

- Issue an index-linked bond.
- Enter into an index-linked swap, with a notional value based on the optimum share of its total funding that should be inflation-linked, in which it pays inflation-linked cash flows and receives fixed-rate income.

The choice will depend on which approach provides cheapest funding and most flexibility.

Hedging a bond issue

Assume that a bank or corporate intends to issue an index-linked bond, and wishes to hedge against a possible fall in government index-linked bond prices, against which its issue will be priced. It can achieve this hedge using an index-linked derivative contract.

The bank or corporate enters into cash-settled contract for difference (CFD), which pays out in the event of a rise in government index-linked bond yields. The CFD has a term to maturity that ties in with the issue date of the index-linked bond. The CFD market maker has effectively shorted the government bond, from the CFD trade date until maturity. On the issue date, the market maker will provide a cash settlement if yields have risen. If yields have fallen, the index-linked bond issuer will pay the difference. However, this cost is netted out by the expected 'profit' from the cheaper funding when the bond is issued. Meanwhile, if yields have risen and the bank or corporate issuer does have to fund at a higher rate, it will be compensated by the funds received by the CFD market maker.

Appendix 10.1: Current issuers of public-sector indexed securities

Country	*Date first issued*	*Index linking*
Australia	1983	Consumer prices
	1991	Average weekly earnings
Austria	1953	Electricity prices
Country	*Date first issued*	*Index linking*
Brazil	1964–90	Wholesale prices
	1991	General prices
Canada	1991	Consumer prices
Chile	1966	Consumer prices
Colombia	1967	Wholesale prices
	1995	Consumer prices
Czech Republic	1997	Consumer prices
Denmark	1982	Consumer prices
France	1956	Average value of French securities
Greece	1997	Consumer prices
Hungary	1995	Consumer prices
Iceland	1964–80	Cost of Building Index
	1980–94	Credit Terms Index
	1995	Consumer prices
Ireland	1983	Consumer prices
Italy	1983	Deflator of GDP at factor cost
Mexico	1989	Consumer prices
New Zealand	1977–84	Consumer prices
	1995	Consumer prices
Norway	1982	Consumer prices
Poland	1992	Consumer prices
Sweden	1952	Consumer prices
	1994	Consumer prices
Turkey	1994–7	Wholesale prices
	1997	Consumer prices
United Kingdom	1981	Consumer prices
United States	1997	Consumer prices

Source: Bloomberg L.P.

Appendix 10.2: US Treasury inflation-indexed securities (TIPS)

Indexation calculation

US TIPS link their coupon and principal to an Index Ratio of the Consumer Prices Index. The index ratio is given by:

$$IR = \frac{CPI_{Settlement}}{CPI_{Issue}}$$

where 'settlement' is the settlement date and 'issue' is the issue date of the bond. The actual CPI used is that recorded for the calendar month three months earlier than the relevant date, this being the lag time. For the first day of any month, the reference CPI level is that recorded three months earlier, so for example on 1 May the relevant CPI measure would be that recorded on 1 February. For any other day in the month, linear interpolation is used to calculate the appropriate CPI level recorded in the reference month and the following month.

Cash flow calculation

The inflation adjustment for the security cash flows is given as the principal multiplied by the index ratio for the relevant date, minus the principal (*P*). This is termed the inflation compensation (*IC*), given as

$$Inflation\ Compensation_{Set\ Date} = (Principal \times Index\ Ratio_{SetDate}) - Principal$$

Coupon payments are given by

$$Interest_{DivDate} = \frac{C}{2} \times (P + IC_{DivDate})$$

The redemption value of a TIPS is guaranteed by the Treasury to be a minimum of US$100 whatever value has been recorded by the CPI during the life of the bond.

Settlement price

The price/yield formula for a TIPS security is given by the following expressions:

$$Price = Inflation - Adjusted\ price + Inflation - Adjusted\ accrued\ interest$$

$$Inflation - Adjusted\ price = Real\ price \times Index\ Ratio_{SetDate}$$

The real price is given by

$$Real\ Price = \left[\frac{1}{1+\frac{f}{d}\frac{r}{2}}\right]\left(\frac{C}{2} + \frac{C}{2}\sum_{j=1}^{n}\phi_j + 100\,\phi^n\right) - RAI \tag{10.18}$$

where

$$\text{Inflation} - \text{Adjusted accrued interest} = RAI \times IR_{SetDate}$$

and where

$$\phi = \left(\frac{1}{1+\frac{r}{2}}\right)$$

r is the annual real yield

RAI is the unadjusted accrued interest, which is $\frac{C}{2} \times \frac{(d-f)}{d}$

f is the number of days from the settlement date to the next coupon date

d is the number of days in the regular semi-annual coupon period ending on the next coupon date

n is the number of full semi-annual coupon periods between the next coupon date and the maturity date.

Selected bibliography and references

Amin, K. and Jarrow, R. 'Pricing foreign currency options under stochastic interest rates', *Journal of International Money and Finance*, 10, 1991, pp. 310–29.

Anderson, N., Breedon, F., Deacon, M., Derry, A. and Murphy, J. *Estimating and Interpreting the Yield Curve*, Wiley, 1996.

Arak, M. and Kreicher, L. 'The real rate of interest: inferences from the new UK indexed gilts', *International Economic Review*, 26, 2, 1985, pp. 399–408.

Bootle, R. *Index-Linked Gilts: A Practical Investment Guide,* 2nd edn, Woodhead-Faulkner, 1991.

Brown, R. and Schaefer, S. 'The term structure of real interest rates and the Cox, Ingersoll and Ross model', *Journal of Financial Economics*, 35, 1, 1994, pp. 3–42.

Brynjolfsson, J. and Fabozzi, F. *Handbook of Inflation Indexed Bonds*, FJF Associates, 1999.

Deacon, M. and Derry, A. 'Deriving estimates of inflation expectations from the prices of UK Government bonds', Bank of England working paper no. 23, July 1994.

Deacon, M. and Derry, A. *Inflation Indexed Securities*, Prentice Hall, 1998.

Foresi, S., Penati, A. and Pennacchi, G. 'Reducing the cost of government debt: the role of index-linked bonds', in de Cecco, M., Pecchi, L. and Piga, G. (eds), *Public Debt: Index-Linked Bonds in Theory and Practice*, Edward Elgar, 1997.

Hughston, L. *Inflation Derivatives*, Merrill Lynch International, 1998.

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

Jarrow, R. and Yildirm, Y. 'Pricing treasury inflation-protected securities and related derivatives', *Journal of Financial and Quantitative Analysis*, 38, 2003.

Van Bezooyen, J., Exley, C. and Smith, A. 'A market-based approach to valuing LPI liabilities', Group for Economic and Market Value based Studies, 1997.

Waggoner, D. 'Spline methods for extracting interest rate curves from coupon bond prices', Federal Reserve Bank of Atlanta working paper, 97–10, 1997.

Wojnilower, A. *Inflation-Indexed Bonds: Promising the Moon*, Clipper Group, 1997.

11
Term Structure Construction Using QuantLib

Term structures are fundamental to the understanding of interest-rate instruments and the pricing of those instruments. It is impossible to understand even simple interest-rate instruments without having at least a rudimentary grasp of what a term structure is and how one is constructed.

In the first edition of this book we explained how to construct term structures aided by a simple piece of software that we built for the purpose, called RATE. However, since writing the first edition there have been huge developments in on-line tools and we believe that it is now more valuable to look at one of these web resources instead, which provide a wealth of solutions in the valuation of financial instruments.

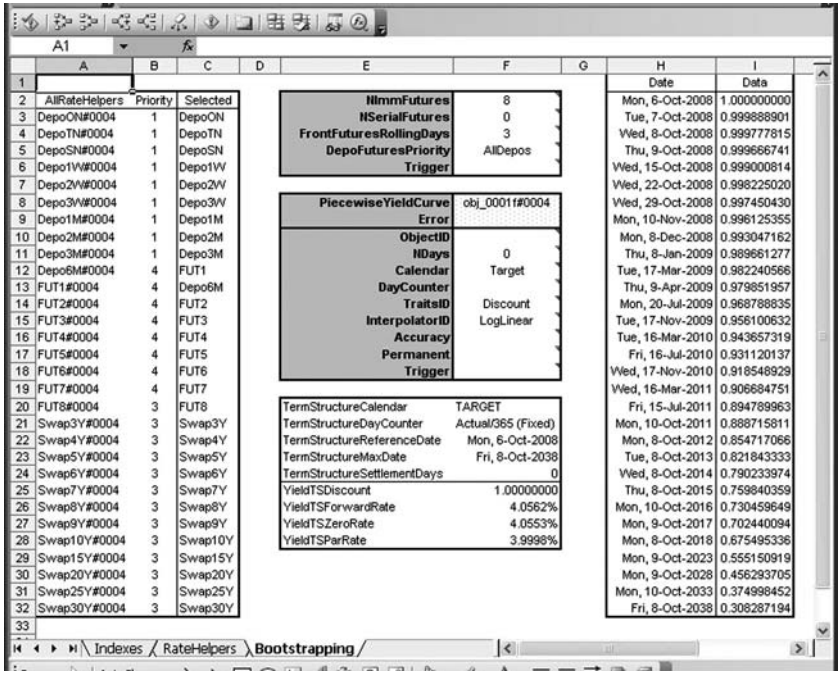
Getting started

We have decided to use www.quantlib.org (QuantLib), or more specifically their Excel add-in, for the purposes of demonstrating yield curve construction. QuantLib is a free library for quantitative finance. There is no need for us to provide a detailed description of what QuantLib does or how to use it because this can all be found on their web site.

What we have done instead is provide additional information on term structure construction using QuantLib to demonstrate the concepts. We downloaded and installed version 0.9.0 of the add-in. For the first step, install QuantLibxl from http://www.quantlibaddin.org.

Open the QuantLib example spreadsheets from the directory where the library was installed. In particular we will use a spreadsheet called 'YieldCurveBootstrapping.xls'. When you open it you will be presented with something that looks like Figure 11.1.

Note that if one opens this on a weekend, the functions may return an error. This can be fixed by using the QuantLib function = *qlSettingsSetEvaluationDate (date, trigger)*. We simply added this function anywhere on the spreadsheet and pointed it at an existing business date and trigger. Of course one could cheat and change the systems date on the computer, but then of course do not forget to change it back when the work is done.



AllRateHelpers	Priority	Selected
DepoON#0004	1	DepoON
DepoTN#0004	1	DepoTN
DepoSN#0004	1	DepoSN
Depo1W#0004	1	Depo1W
Depo2W#0004	1	Depo2W
Depo3W#0004	1	Depo3W
Depo1M#0004	1	Depo1M
Depo2M#0004	1	Depo2M
Depo3M#0004	1	Depo3M
Depo6M#0004	4	FUT1
FUT1#0004	4	Depo6M
FUT2#0004	4	FUT2
FUT3#0004	4	FUT3
FUT4#0004	4	FUT4
FUT5#0004	4	FUT5
FUT6#0004	4	FUT6
FUT7#0004	4	FUT7
FUT8#0004	3	FUT8
Swap3Y#0004	3	Swap3Y
Swap4Y#0004	3	Swap4Y
Swap5Y#0004	3	Swap5Y
Swap6Y#0004	3	Swap6Y
Swap7Y#0004	3	Swap7Y
Swap8Y#0004	3	Swap8Y
Swap9Y#0004	3	Swap9Y
Swap10Y#0004	3	Swap10Y
Swap15Y#0004	3	Swap15Y
Swap20Y#0004	3	Swap20Y
Swap25Y#0004	3	Swap25Y
Swap30Y#0004	3	Swap30Y

NImmFutures	8
NSerialFutures	0
FrontFuturesRollingDays	3
DepoFuturesPriority	AllDepos
Trigger	

PiecewiseYieldCurve	obj_0001f#0004
Error	
ObjectID	
NDays	0
Calendar	Target
DayCounter	
TraitsID	Discount
InterpolatorID	LogLinear
Accuracy	
Permanent	
Trigger	

TermStructureCalendar	TARGET
TermStructureDayCounter	Actual/365 (Fixed)
TermStructureReferenceDate	Mon, 6-Oct-2008
TermStructureMaxDate	Fri, 8-Oct-2038
TermStructureSettlementDays	0
YieldTSDiscount	1.00000000
YieldTSForwardRate	4.0562%
YieldTSZeroRate	4.0553%
YieldTSParRate	3.9998%

Date	Data
Mon, 6-Oct-2008	1.000000000
Tue, 7-Oct-2008	0.999888901
Wed, 8-Oct-2008	0.999777815
Thu, 9-Oct-2008	0.999666741
Wed, 15-Oct-2008	0.999000814
Wed, 22-Oct-2008	0.998225020
Wed, 29-Oct-2008	0.997450430
Mon, 10-Nov-2008	0.996125355
Mon, 8-Dec-2008	0.993047162
Thu, 8-Jan-2009	0.989661277
Tue, 17-Mar-2009	0.982240586
Thu, 9-Apr-2009	0.979851957
Mon, 20-Jul-2009	0.968788835
Tue, 17-Nov-2009	0.956100632
Tue, 16-Mar-2010	0.943657319
Fri, 16-Jul-2010	0.931120137
Wed, 17-Nov-2010	0.918548929
Wed, 16-Mar-2011	0.906684751
Fri, 15-Jul-2011	0.894789963
Mon, 10-Oct-2011	0.888715811
Mon, 8-Oct-2012	0.854717066
Tue, 8-Oct-2013	0.821843333
Wed, 8-Oct-2014	0.790233974
Thu, 8-Oct-2015	0.759840359
Mon, 10-Oct-2016	0.730459649
Mon, 9-Oct-2017	0.702440094
Mon, 8-Oct-2018	0.675495336
Mon, 9-Oct-2023	0.555150919
Mon, 9-Oct-2028	0.456293705
Mon, 10-Oct-2033	0.374998452
Fri, 8-Oct-2038	0.308287194

Figure 11.1 Yield curve screen

Using QuantLib's yield curve bootstrapping

Capturing your market data

QuantLib's yield curve model constructs a discount, zero yield or forward rate curve using a combination of short-term money market, traded futures and OTC swap instruments. The yield curve screen does not contain the market data. The market data is held in a separate tab in the spreadsheet called 'RateHelpers'. It is shown at Figure 11.2.

A zero curve is defined by the value date and QuantLib version 0.9.0 assumes that one is building a zero curve from the current date.

Selecting your calculation parameters

In order to construct the curve, QuantLib requires calculation parameters. If calculation parameters are not provided then QuantLib will use defaults. The parameters required are holiday calendar, interpolation method and day count basis. QuantLib supports a number of parameters and has pre-existing data for certain calendars. Figure 11.3 is an illustration of the day-counter setting.

Various types of yield curve interpolation are supported by QuantLib including linear, log-linear and cubic-spline. A selection of interpolation methods are discussed later in the chapter.

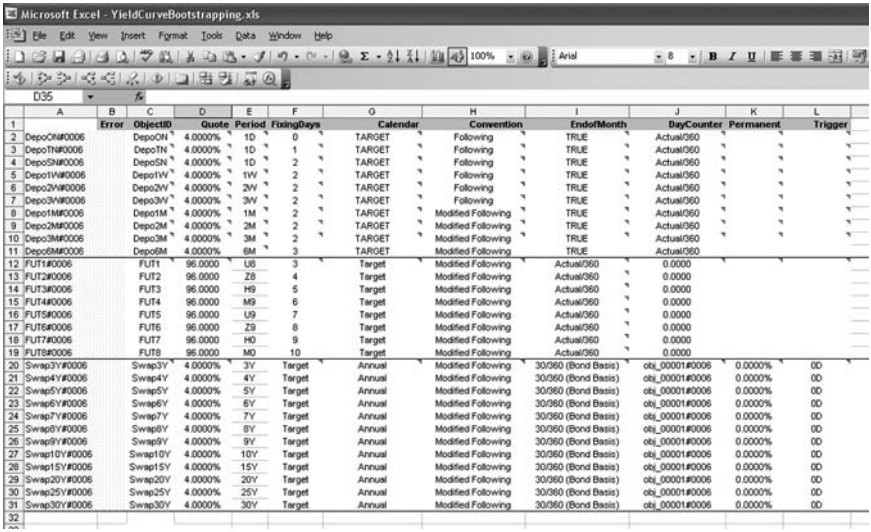


Figure 11.2 Market data screen

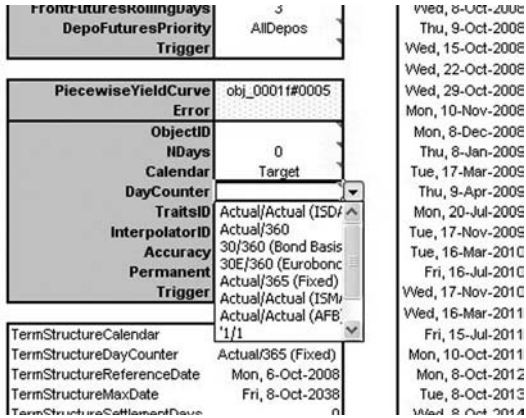


Figure 11.3 Day counter setting

To calculate a curve, QuantLib needs to know what output options to apply. QuantLib displays discount factors and also allows the user to convert discount factors into zero rates or forward rates. The illustration of the options for setting output is shown at Figure 11.4.

Standard yield curve construction methodology

There is a standard methodology to building yield curves which must start from the earliest date and work towards the later dates. Firstly the money market curve is built, then the futures are spliced with the money market, and finally the swaps are bootstrapped.

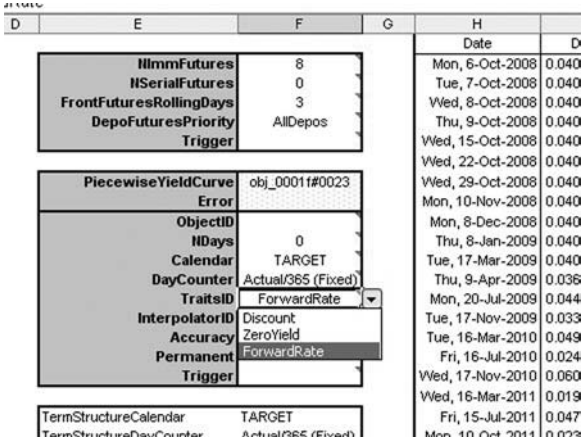


Figure 11.4 Output setting

The money market discount factors are calculated using each money market quote from the earliest maturity date to the first maturity date that overlaps the swap inputs.

Where money market inputs mature in less than one year, a discount function that has no compounding is applied. Money market discount factors for periods of more than one year are calculated using a discount function that has annual compounding. These equations are set out below. When calculating discount factors for a maturity that is greater than one year we use:

$$df = \frac{1}{(1+r)^B} \tag{11.1}$$

When calculating discount factors for a maturity that is less than one year we use:

$$df = \frac{1}{1+(r \times B)} \tag{11.2}$$

where B is the day count fraction and r is the quoted money market rate. The day count where fraction is discussed later in this chapter. Note that it is the day count of the money market instrument being used as an input which is relevant, not the day count used to stipulate the yield curve outputs.

We cannot ignore the fix days in the calculation. For the first money market deposit the fixing is zero days so we can simply apply formula (11.2). However, the fix days for the 'tom-next' deposit is one day. This is therefore the quoted rate for a deposit starting in one day for the term of just one night. So once we have calculated the discount factor for one night we need to multiply it by the discount factor of the first day on our term structure as follows:

$$df_1 = \frac{1}{1+(r \times B)} \times df_0. \tag{11.3}$$

Similar adjustments are required for subsequent deposits which have two (or three) fix days. Once all money market quotes are translated into discount factors, the futures contract prices are spliced with the curve. For the purposes of date generation, futures normally have contracts with start dates on the third Wednesday of each quarter and settlement is two days after expiry. Each contract is assumed to be exactly 30 days multiplied by the monthly length of the contract. QuantLib tells you whether the deposit rate or future has been given priority in the calculation.

On the assumption that futures contracts are less than 360 days in length, their discount factors are calculated using a function that has no compounding:

$$df_{90} = \frac{1}{1 + ((\frac{100 - P_{fut}}{100}) \times B_{90})} \tag{11.4}$$

where df_{90} is the discount factor for the 90 day futures contract period. This is really a 92-day period because of the 2-day delay in settlement. P_{fut} is the future contract quoted price and B_{90} is the day count fraction for the future (a 92-day period).

Any discount factor calculation requires two dates, the start date and the maturity date. The start date for money market quotes is a date that coincides with or is close to the starting point of the curve (that is, the value date). Futures, however, have a future start date. In other words, when we calculate the discount factor for a futures contract, it does not represent a calculation for today but rather a discount factor for a strip of time (90 days) starting on a future date. This means that this discount factor needs to be spliced with one that runs from the value date to the start of the futures contract. This is illustrated at Figure 11.5.

Displayed at Figure 11.5 is a futures contract F_1. One can calculate the discount factor for the contract F_1 using the formula (11.3). However, this is only applicable to its future strip, the 90 day time period F_0 to F_1. F_1's start date falls between two money market contracts M_3 and M_4. In order to get the discount factor applicable

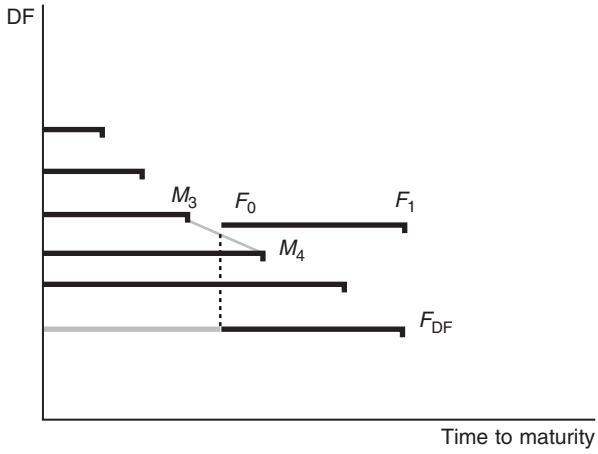


Figure 11.5 Futures splicing

at the start of the contract, interpolate between M_3 and M_4. The futures contract F_1 can now be spliced with the curve by multiplying the discount factor calculated by interpolation with the discount factor for F_1. This provides F_{DF}'s discount factor, a the discount factor for the entire period. Each and every subsequent futures contract can then be spliced in the same way by applying this formula.

$$df_T = \frac{1}{1+((\frac{100-P_{fut}}{100})\times B_{90})} \times df_m \tag{11.5}$$

In equation (11.5) df_T is the discount factor for the period from value date to the futures contract maturity date and df_m is the money market discount factor at the futures contract start date, calculated by interpolation.

Bootstrapping

Interest rate swap discount factors are calculated using bootstrapping. Bootstrapping also applies to bonds and so this section is equally applicable to a yield curve constructed using bond quotes. Unlike the cash and futures contracts, bonds and swaps have periodic interest payments or coupons. The quoted rate for a swap contract, or yield to maturity on a bond, represents not only the settlement at maturity date but also the collective rate for each and every interest coupon. The quoted rate represents the cumulative quote of a number of rates that span the term structure of the bond or swap contract.

The construction of a zero curve needs to strip out these various coupon payments and identify the appropriate rate applied to each. This is done by 'bootstrapping' the swap or bond.

In Figure 11.6 a swap contract with settlement dates i_1, i_2, i_3 and i_4 is shown. In addition, money market contracts M_3 and M_4 and futures contract F_{DF}, for which

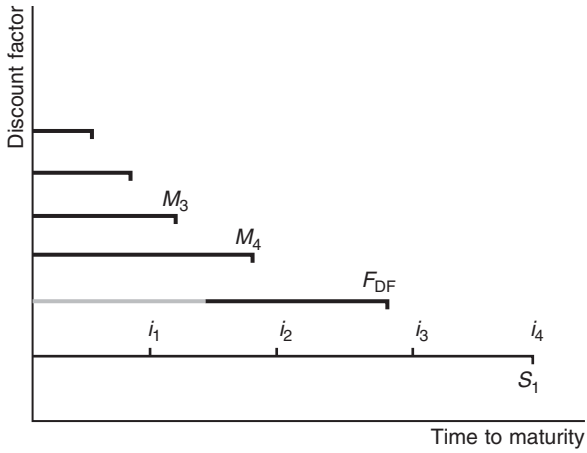


Figure 11.6 Bootstrapping

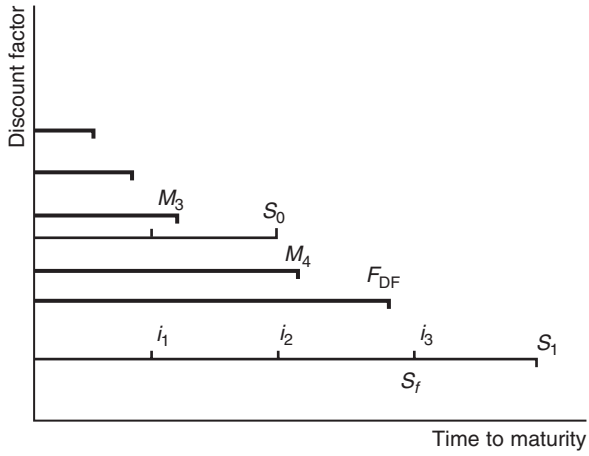


Figure 11.7 Overlapping coupon

discount factors have been calculated, are available. These discount factors, together with a choice of interpolation, can now be used to calculate an appropriate discount factor for coupon points i_1 and i_2.

The difficulty with i_3 is that this swap coupon has a settlement date that lies past the previously calculated discount factors. Interpolation cannot be used to calculate a discount factor for i_3. In order to solve this problem we can create a fictitious swap. This is shown at Figure 11.7.

This fictitious instrument, S_f, is a sub set of S_1. It is the same as S_1 up to and including coupon i_3 but excludes the last coupon. In other words, the fictitious maturity date is point i_3 and not point i_4.

Our intention is to calculate a discount factor for S_f. However, a discount factor for S_f cannot be calculated because we do not have a swap rate. We cannot simply use the swap rate for swap contract S_1 because S_1 has a different maturity and its rate would be inappropriate. A rate for S_f needs to be estimated using interpolation. The swap rate for the previous swap, S_0, and the next swap, which is now S_1, are used. By using interpolation a rate for S_f is calculated from the quoted rate of S_0 and S_1.

Whether we create a fictitious swap or not, we are now in a position to calculate the zero-coupon discount factor for the swap S_f. The function that is applied is:

$$df = \frac{1-(r\times\sum_{i=1}^{t-1}(df_i\times B_i))}{1+(r\times B_t)} \tag{11.6}$$

where r is the quoted swap rate, and df_i is the interpolated discount factor for the end of each coupon period excluding the last coupon settled at maturity. Similarly, B_I is the day count fraction for each coupon period from 1 to $t-1$

(that is, all excluding the last coupon), and B_t is the day count fraction for the last coupon period.

Each swap discount factor can now be calculated in turn. S_1 once calculated can be fed into the calculation of S_2, and S_2 into S_3, and so on. This is what is termed 'bootstrapping'. In other words, when one laces one's bootstraps one must start from the bottom and work one's way up to the top.

Bonds can be bootstrapped in much the same way. There are two key differences when applying bond bootstrapping. Firstly, there are no cash or futures markets in the bond model. The first bond is therefore not bootstrapped. Instead, the yield to maturity for the first bond is applied to all coupon and principal settlement points on the bond. Secondly, a bond has a price. This is applied to the bootstrapping function as follows:

$$df = \frac{P_{bond} - (1 \times r \times \sum_{i=1}^{t-1} (df_i \times B_i))}{1 + (1 \times r \times B_t)}. \tag{11.7}$$

Interpolation

The cash, futures and swaps quotes are quotes for a set of fixed points in time (tenors). When we need a rate or discount factor for a point in time that does not coincide exactly with the quoted tenors, a method of estimating the rate or discount factor is required. This is achieved using interpolation.

There are various methods of interpolation that have become widely accepted and used in the market; for example, linear, exponential and cubic-spline interpolation.

In Figure 11.8 discount factors have been calculated for a number of points on the zero curve. How does one determine the discount factor for a date D_3 that lies between points P_1 and P_2?

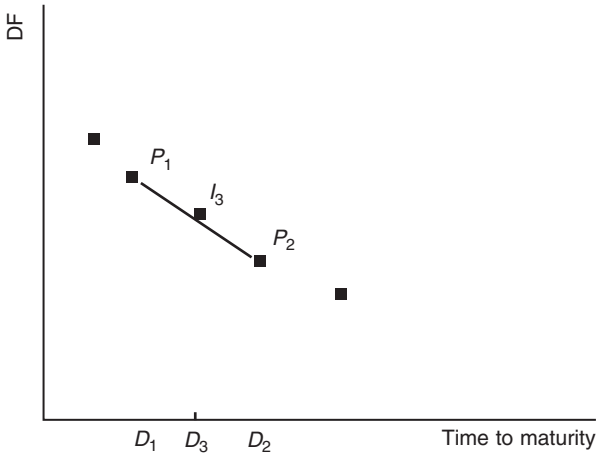


Figure 11.8 Interpolation

Using linear interpolation a straight line is drawn between the two points P_1 and P_2 and where they cross the date D_3, the interpolated rate is determined. This is achieved mathematically using the following function:

$$I_3 = P_1 + ((\tfrac{P_2 - P_1}{D_2 - D_1}) \times (D_3 - D_1)) \tag{11.8}$$

The determination of exponential interpolation uses the following calculation:

$$I_3 = P_1^{\partial_1} \times P_2^{\partial_2} \tag{11.9}$$

where:

$$\partial_1 = \frac{D_3 - D_0}{D_1 - D_0} \times T;$$

$$\partial_2 = \frac{D_3 - D_0}{D_2 - D_0} \times (1 - T); \text{ and}$$

$$T = \frac{D_2 - D_3}{D_2 - D_1}$$

Business days

Contracts settle on business days. If a payment date happens to fall on a weekend or public holiday, then the nearest business day needs to be determined. Two possible business day conventions are:

1. Following business day
2. Preceding business day.

The following business day identifies the first business day that falls after the calculated date. Similarly, the preceding business day looks for the first business day prior to the calculated date. Both of these methods can be modified. In the case of the modified rule, the following business day is applied, but if the following business day falls in the next month, then the preceding business day basis is used. Similarly, the modified preceding method changes to the following business day method if the preceding method results in a date that is in the prior month.

Day count basis and the day count fraction

Calculation of future value (with annual compounding) is given by:

$$FV = PV \times (1 + r)^B. \tag{11.10}$$

The day count fraction is a consistent means by which B can be reflected after taking into consideration the day count basis; for example, actual/360 or actual/365.

For example using a day count basis of actual/365 the day count fraction would be calculated using the following formula;

$$B = \frac{d}{365} \tag{11.11}$$

where d is the number of days from value date to maturity or the contract length expressed in days. There are a number of business day conventions, each with their own quirks which can be applied when using the QuantLib libraries.

We recommand that readers explore the QuantLib Excel add-in and apply it to constructing yield curves for their own particular market.

This page intentionally left blank

PART III
Structured Financial Products

The market in structured financial products is very large and diverse. In Part III of the book we aim to give readers a flavour of these instruments. After introducing the concept of securitisation, we discuss mortgage-backed securities and collateralised debt obligations or CDOs. We also discuss synthetic securitised products.

This market was severely impacted by the 2007–2008 financial crash (see chapter 24) and many types of structured finance products disappeared. However asset-backed securities, and the technique of securitisation, remain important in finance.

This page intentionally left blank

12
An Introduction to Asset-Backed Securities and Securitisation[1]

In this chapter we introduce the basic concepts of securitisation and look at the motivation behind their use, as well as their economic impact. We illustrate the process with a brief hypothetical case study. We also describe the 'in-house' deal, a response by banks to the 2007–8 credit crunch.

The concept of securitisation

Securitisation is a well-established practice in the global debt capital markets. It refers to the sale of assets, which generate cash flows, from the institution that owns them, to another company that has been specifically set up for the purpose, and the issuing of notes by this second company. These notes are backed by the cash flows from the original assets. The technique was introduced initially as a means of funding for US mortgage banks. Subsequently the technique was applied to other assets such as credit card payments and leasing receivables. It has also been employed as part of asset/liability management, as a means of managing balance sheet risk.

Securitisation allows institutions such as banks and corporates to convert assets that are not readily marketable – such as residential mortgages or car loans – into rated securities that are tradeable in the secondary market. The investors that buy these securities gain an exposure to these types of original assets, to which they would not otherwise have access. The technique is well established and was first introduced by mortgage banks in the United States in 1979. The later synthetic securitisation market is much more recent, dating from 1997. The key difference between cash and synthetic securitisation is that in the former, as we have noted, the assets in question are actually sold to a separate legal company known as a special-purpose vehicle (SPV, also referred to as a special-purpose entity (SPE) or special-purpose company (SPC)). This does not occur in a synthetic transaction, as we shall see.

Sundaresan defines securitisation as 'a framework in which some illiquid assets of a corporation or a financial institution are transformed into a package of

[1]This chapter was co-authored with Anuk Teasdale of YieldCurve.com.

securities backed by these assets, through careful packaging, credit enhancements, liquidity enhancements and structuring' (1997, p. 359).

The process of securitisation creates *asset-backed bonds*. These are debt instruments that have been created from a package of loan assets on which interest is payable, usually on a floating basis. The asset-backed market was developed in the United States and is a large, diverse market containing a wide range of instruments. Techniques employed by investment banks today enable an entity to create a bond structure from any type of cash flow; assets that have been securitised include loans such as residential mortgages, car loans and credit card loans. The loans form assets on a bank or finance house balance sheet, which are packaged together and used as backing for an issue of bonds. The interest payments on the original loans form the cash flows used to service the new bond issue. Traditionally mortgage-backed bonds are grouped in their own right as *mortgage-backed securities* (MBS) while all other securitisation issues are known as *asset-backed bonds* or ABS.

Reasons for undertaking securitisation

The driving force behind securitisation has been the need for banks to realise value from the assets on their balance sheet. Typically these assets are residential mortgages, corporate loans, and retail loans such as credit card debt. Let us consider the factors that might lead a financial institution to securitise a part of its balance sheet. These might be for the following reasons:

- If revenues received from assets remain roughly unchanged but the size of assets has decreased, this will lead to an increase in the return on equity ratio.
- The level of capital required to support the balance sheet will be reduced, which again can lead to cost savings or allows the institution to allocate the capital to other, perhaps more profitable, business.
- To obtain cheaper funding: frequently the interest payable on ABS securities is considerably below the level payable on the underlying loans. This creates a cash surplus for the originating entity.

In other words a bank will securitise part of its balance sheet for one or all of the following main reasons:

- funding the assets it owns
- balance sheet capital management
- risk management and credit risk transfer.

We consider each of these in turn.

Funding

Banks can use securitisation to support rapid asset growth, diversify their funding mix and reduce the cost of funding, and reduce maturity mismatches. Up until the 2007–2008 crash, the market for asset-backed securities was large, with an estimated

amount of US$1,000 billion invested in ABS issues worldwide annually, of which US$150 billion is in the European market alone.

Securitising assets also allows a bank to diversify its funding mix. All banks will not wish to be reliant on only a single or a few sources of funding, as this can be high-risk in times of market difficulty. Banks aim to optimise their funding between a mix of retail, interbank and wholesale sources. Securitisation has a key role to play in this mix. It also enables a bank to reduce its funding costs. This is because the securitisation process de-links the credit rating of the originating institution from the credit rating of the issued notes. Typically most of the notes issued by SPVs will be higher rated than the bonds issued directly by the originating bank. While the liquidity of the secondary market in ABS is frequently lower than that of the corporate bond market, and this adds to the yield payable by an ABS, it is frequently the case that the cost to the originating institution of issuing debt is still lower in the ABS market because of the latter's higher rating.

Finally, there is the issue of maturity mismatches. The business of bank ALM is inherently one of maturity mismatch, since a bank often funds long-term assets such as residential mortgages with short-asset liabilities such as bank account deposits or interbank funding. This mismatch can be removed via securitisation, as the originating bank receives funding from the sale of the assets, and the economic maturity of the issued notes frequently matches that of the assets.

Balance sheet capital management

Banks use securitisation to improve balance sheet capital management. This provides regulatory capital relief, economic capital relief and diversified sources of capital. As stipulated in the Bank for International Settlements (BIS) capital rules,[2] also known as the Basel rules, banks must maintain a minimum capital level for their assets, in relation to the risk of these assets. Under Basel I, for every US$100 of risk-weighted assets a bank must hold at least US$8 of capital; however the designation of each asset's risk-weighting is restrictive. For example with the exception of mortgages, customer loans are 100% risk-weighted regardless of the underlying rating of the borrower or the quality of the security held. The anomalies that this raises, which need not concern us here, were partly addressed by the Basel II rules which became effective in 2007. However the Basel I rules were a key driver of securitisation.[3]

As an SPV is not a bank, it is not subject to Basel rules and needs only such capital as is economically required by the nature of the assets it contains. This is not a set amount, but is significantly below the 8% level required by banks in all cases. Although an originating bank does not obtain 100% regulatory capital relief when it sells assets off its balance sheet to an SPV, because it will have retained a 'first-loss' piece out of the issued notes, its regulatory capital charge will be significantly reduced after the securitisation.[4]

[2] For further information on this see Choudhry (2001).
[3] For further detail on the Basel rules see Choudhry (2007) and Choudhry (2010).
[4] We discuss first loss in Chapter 14.

To the extent that securitisation provides regulatory capital relief, it can be thought of as an alternative to capital raising, compared with the traditional sources of Tier 1 capital (equity), preferred shares, and perpetual loan notes with step-up coupon features. By reducing the amount of capital that has to be used to support the asset pool, a bank can also improve its return-on-equity (ROE) value. This will be received favourably by shareholders.

Risk management

Once assets have been securitised, the credit risk exposure on these assets for the originating bank is reduced considerably, and if the bank does not retain a first-loss capital piece (the most junior of the issued notes), it is removed entirely. This is because assets have been sold to the SPV. Securitisation can also be used to remove non-performing assets from banks' balance sheets. This has the dual advantage of removing credit risk and removing a potentially negative sentiment from the balance sheet, as well as freeing up regulatory capital as before. Further, there is a potential upside from securitising such assets: if any of them start performing again, or there is a recovery value obtained from defaulted assets, the originator will receive any surplus profit made by the SPV.

Benefits of securitisation to investors

Investor interest in the ABS market has been considerable from the market's inception. This is because investors perceive ABSs as possessing a number of benefits. Investors can:

- diversify sectors of interest
- access different (and sometimes superior) risk–reward profiles
- access sectors that are otherwise not open to them.

A key benefit of securitisation notes is the ability to tailor risk–return profiles. For example, if there is a lack of assets of any specific credit rating, these can be created via securitisation. Securitised notes frequently offer better risk–reward performance than corporate bonds of the same rating and maturity. While this might seem peculiar (why should one AA-rated bond perform better in terms of credit performance than another just because it is asset-backed?), this often occurs because the originator holds the first-loss piece in the structure.

The process of securitisation

We look now at the process of securitisation, the nature of the SPV structure, and issues such as credit enhancements and the cash flow waterfall.

The securitisation process

The securitisation process involves a number of participants. In the first instance there is the originator, the firm whose assets are being securitised. The most

common process involves an issuer acquiring the assets from the *originator*. The issuer is usually a company that has been specially set up for the purpose of the securitisation, which is the SPV and is usually domiciled offshore. The creation of an SPV ensures that the underlying asset pool is held separate from the other assets of the originator. This is done so that in the event that the originator is declared bankrupt or insolvent, the assets that have been transferred to the SPV will not be affected. This is known as being bankruptcy-remote. Conversely, if the underlying assets begin to deteriorate in quality and are subject to a ratings downgrade, investors have no recourse to the originator.

When the assets are held within an SPV framework, defined in formal legal terms, the financial status and credit rating of the originator becomes almost irrelevant to the bondholders. The process of securitisation often involves *credit enhancements*, in which a third-party guarantee of credit quality is obtained, so that notes issued under the securitisation are often rated at investment grade and up to AAA-grade.

The process of structuring a securitisation deal ensures that the liability side of the SPV – the issued notes – carries lower cost than the asset side of the SPV. This enables the originator to secure lower-cost funding that it would not otherwise be able to obtain in the unsecured market. This is a tremendous benefit for institutions with lower credit ratings. Figure 12.1 illustrates the process of securitisation in simple fashion.

Mechanics of securitisation

Securitisation involves a 'true sale' of the underlying assets from the balance sheet of the originator. This is why a separate legal entity, the SPV, is created to act as the issuer of the notes. The assets being securitised are sold onto the balance sheet of

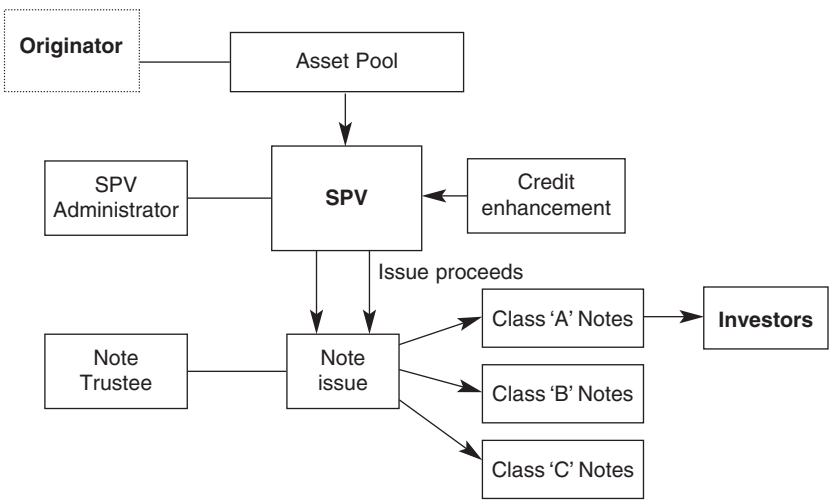


Figure 12.1 The process of securitisation

the SPV and are, therefore, ring-fenced from those of the originating institution. The process involves:

- undertaking 'due diligence' on the quality and future prospects of the assets
- setting up the SPV and then effecting the transfer of assets to it
- underwriting of loans for credit quality and servicing
- determining the structure of the notes, including how many tranches are to be issued, in accordance to originator and investor requirements
- the notes being rated by one or more credit rating agencies
- placing of notes in the capital markets.

The sale of assets to the SPV needs to be undertaken so that it is recognised as a true legal transfer. The originator will need to hire legal counsel to advise it in such matters. The credit rating process will consider the character and quality of the assets, and also whether any enhancements have been made to the assets that will raise their credit quality. This can include *overcollateralisation*, which is when the principal value of notes issued is lower than the principal value of assets, and a liquidity facility provided by a bank.

A key consideration for the originator is the choice of the underwriting bank, which structures the deal and places the notes. The originator will award the mandate for its deal to an investment bank on the basis of fee levels, marketing ability and track record with assets being securitised.

SPV structures

There are essentially two main securitisation structures, *amortising* (pass-through) and *revolving*. A third type, the *master trust*, is used by frequent issuers.

Amortising structures

Amortising structures pay principal and interest to investors on a coupon-by-coupon basis throughout the life of the security, as illustrated in Figure 12.2. They are priced and traded based on expected maturity and weighted-average lifer (WAL), which is the time-weighted period during which principal is outstanding. A WAL approach incorporates various pre-payment assumptions, and any change in this pre-payment speed will increase or decrease the rate at which principal is repaid to investors. Pass-through structures are commonly used in residential and commercial mortgage-backed deals (MBS), and consumer loan ABS.

Revolving structures

Revolving structures revolve the principal of the assets; that is, during the revolving period, principal collections are used to purchase new receivables which fulfil the necessary criteria. The structure is used for short-dated assets with a relatively high pre-payment speed, such as credit card debt and car loans. During the amortisation period, principal payments are paid to investors either in a series of equal instalments (*controlled amortisation*) or principal is 'trapped' in a separate account

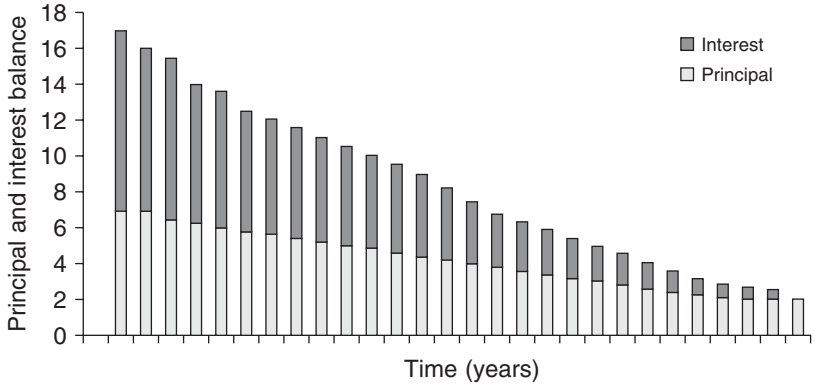


Figure 12.2 Amortising structures

until the expected maturity date and then paid in a single lump sum to investors (*soft bullet*).

Master trust

Frequent issuers under US and UK law use master trust structures, which allow multiple securitisations to be issued from the same SPV. Under such schemes, the originator transfers assets to the master trust SPV. Notes are then issued out of the asset pool based on investor demand. Master trusts have been used by MBS and credit card ABS originators.

Securitisation note tranching

As illustrated in Figure 12.1, in a securitisation the issued notes are structured to reflect specified risk areas of the asset pool, and thus are rated differently. The senior tranche is usually rated AAA. The lower-rated notes usually have an element of *over-collateralisation* and are thus capable of absorbing losses. The most junior note is the lowest-rated or non-rated. It is often referred to as the *first-loss piece*, because it is impacted by losses in the underlying asset pool first. The first-loss piece is sometimes called the *equity piece* or equity note (even though it is a bond) and is usually held by the originator.

Credit enhancement

Credit enhancement refers to the group of measures that can be instituted as part of the securitisation process for ABS and MBS issues so that the credit rating of the issued notes meets investor requirements. The lower the quality of the assets being securitised, the greater the need for credit enhancement. This is usually done by any of the following methods:

- *Over-collateralisation*: where the nominal value of the assets in the pool is in excess of the nominal value of issued securities.

- *Pool insurance*: an insurance policy provided by a composite insurance company to cover the risk of principal loss in the collateral pool. The claims paying rating of the insurance company is important in determining the overall rating of the issue.
- *Senior/junior note classes*: credit enhancement is provided by subordinating a class of notes ('class B' notes) to the senior class notes ('class A' notes). The class B note's right to its proportional share of cash flows is subordinated to the rights of the senior noteholders. Class B notes do not receive payments of principal until certain rating agency requirements have been met, specifically satisfactory performance of the collateral pool over a predetermined period, or in many cases until all of the senior note classes have been redeemed in full.
- *Margin step-up*: a number of ABS issues incorporate a step-up feature in the coupon structure, which typically coincides with a call date. Although the issuer is usually under no obligation to redeem the notes at this point, the step-up feature was introduced as an added incentive for investors, to convince them from the outset that the economic cost of paying a higher coupon would be unacceptable and that the issuer would seek to refinance by exercising its call option.
- *Excess spread*: this is the difference between the return on the underlying assets and the interest rate payable on the issued notes (liabilities). The monthly excess spread is used to cover expenses and any losses. If any surplus is left over, it is held in a reserve account to cover against future losses or (if not required for that), as a benefit to the originator. In the meantime the reserve account is a credit enhancement for investors.

All securitisation structures incorporate a *cash waterfall* process, whereby all the cash that is generated by the asset pool is paid in order of payment priority. Only when senior obligations have been met can more junior obligations be paid. An independent third-party agent is usually employed to run 'tests' on the vehicle to confirm that there is sufficient cash available to pay all obligations. If a test is failed, then the vehicle will start to pay off the notes, starting from the senior notes. The waterfall process is illustrated in Figure 12.3.

Impact on balance sheet

Figure 12.4 illustrates by a hypothetical example the effect on the liability side of an originating bank's balance sheet of a securitisation transaction. Following the process, selected assets have been removed from the balance sheet, although the originating bank will usually have retained the first-loss piece. With regard to the regulatory capital impact, this first-loss amount is deducted from the bank's total capital position. For example, assume a bank has US$100 million of risk-weighted assets and a target Basel ratio of 12%,[5] and securitises all US$100 million of these assets. It retains the first-loss tranche which forms 1.5% of the total issue. The remaining 98.5% will be sold on to the market. The bank will still have to set aside

[5] The minimum is 8% but many banks prefer to set aside an amount well in excess of this minimum required level.

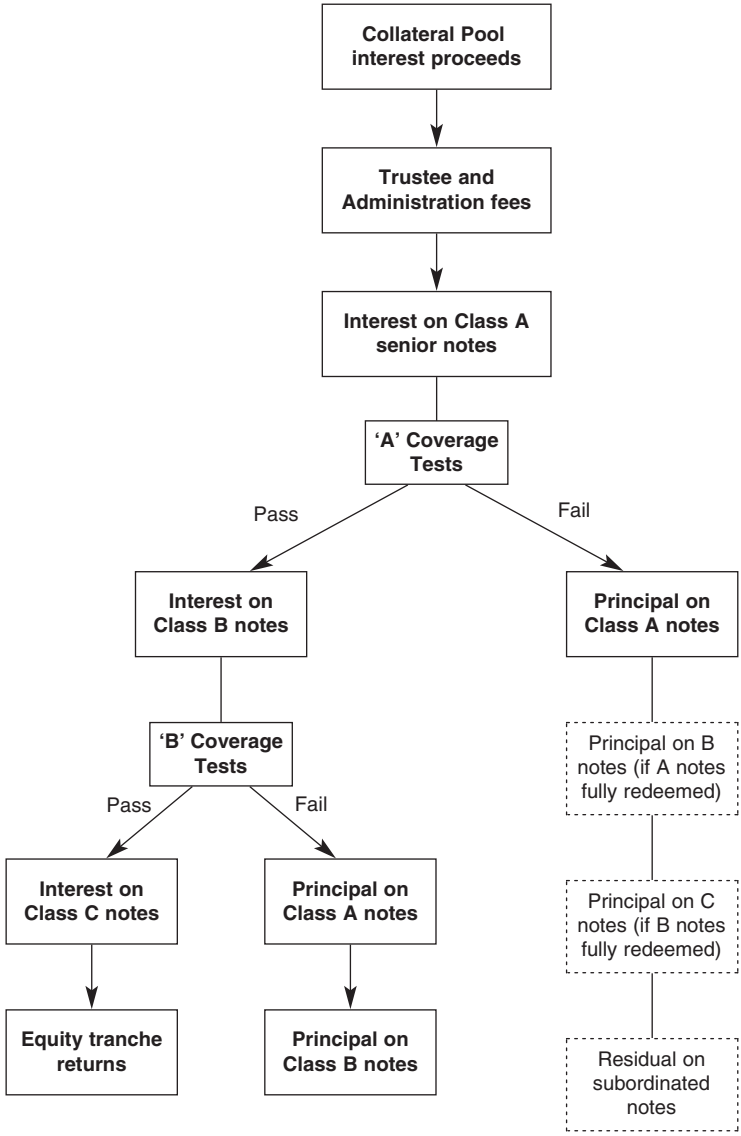


Figure 12.3 The waterfall process

1.5% of capital as a buffer against future losses, but it has been able to free itself of the remaining 10.5% of capital.

Illustrating the process of securitisation

To illustrate the process of securitisation, we consider a hypothetical airline ticket receivables transaction, being originated by the fictitious ABC Airways plc and

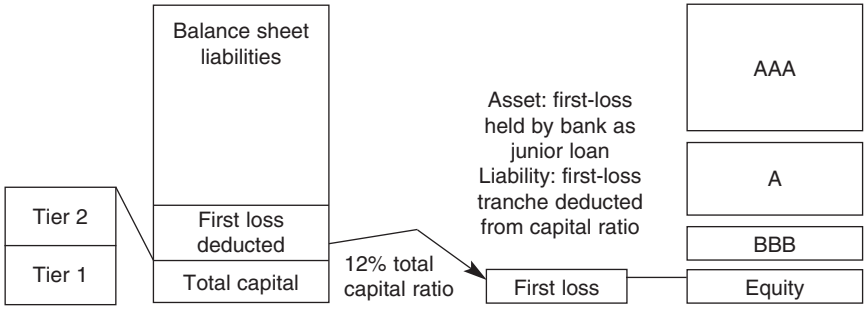


Figure 12.4 The effect on the liability side of the balance sheet of a securitisation transaction

arranged by the equally fictitious XYZ Securities Limited. We show the kind of issues that will be considered by the investment bank that is structuring the deal.

Originator	ABC Airways plc
Issuer	'Airways No 1 Ltd'
Transaction	Ticket receivables airline future flow securitisation bonds
	200 m three-tranche floating rate notes, legal maturity 2010
	Average life 4.1 years
Tranches	Class 'A' note (AA), Libor plus [ ] bps
	Class 'B' note (A), Libor plus [ ] bps
	Class 'E' note (BBB), Libor plus [ ] bps
Arranger	XYZ Securities plc

Due diligence

XYZ Securities will undertake due diligence on the assets to be securitised. For this case, it will examine the airline performance figures over the last five years, as well as model future projected figures, including:

- total passenger sales
- total ticket sales
- total credit card receivables
- geographical split of ticket sales.

It is the future flow of receivables, in this case credit card purchases of airline tickets, that is being securitised. This is a higher-risk asset class than say, residential mortgages, because the airline industry has a tradition of greater volatility of earnings than, say, mortgage banks.

Marketing approach

The present and all future credit card ticket receivables generated by the airline will be transferred to an SPV. The investment bank's syndication desk will seek to

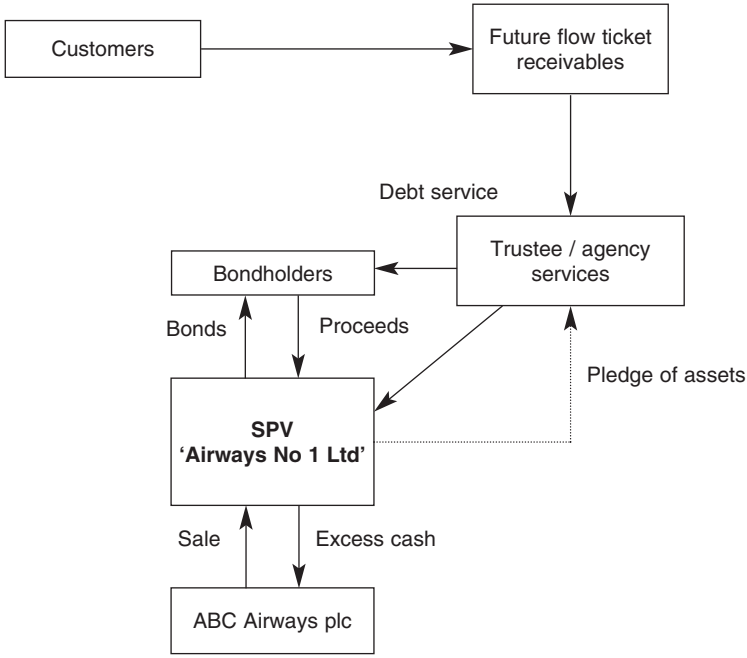


Figure 12.5 Illustrative deal structure

place the notes with institutional investors across Europe. The notes are first given an indicative pricing ahead of the issue, to gauge investor sentiment. Given the nature of the asset class, during November 2002 the notes would be marketed at around three-month Libor plus 70–80 bps (AA note), 120–130 bps (A note) and 260–270 bps (BBB note). The notes are 'benchmarked' against recent issues with similar asset classes, as well as the spread level in the unsecured market of comparable issuer names.

Deal structure

The deal structure is shown at Figure 12.5. The process leading to issue of notes is as follows:

- ABC Airways plc sells its future flow ticket receivables to an offshore SPV set up for this deal, incorporated as Airways No 1 Ltd.
- The SPV issues notes in order to fund its purchase of the receivables.
- The SPV pledges its right to the receivables to a fiduciary agent, the Security Trustee, for the benefit of the bondholders.
- The Trustee accumulates funds as they are received by the SPV.
- The bondholders receive interest and principal payments, in the order of priority of the notes, on a quarterly basis.

In the event of default, the Trustee will act on behalf of the bondholders to safeguard their interests.

Financial guarantors

The investment bank will consider whether an insurance company, known as a monoline insurer, should be approached to 'wrap' the deal by providing a guarantee of backing for the SPV in the event of default. This insurance is provided in return for a fee.

Financial modelling

XYZ Securities will construct a cash flow model to estimate the size of the issued notes. The model will consider historical sales values, any seasonal factors in sales, credit card cash flows, and so on. Certain assumptions will be made when constructing the model, for example growth projections, inflation levels and tax levels. The model will consider a number of different scenarios, and also calculate the minimum asset coverage levels required to service the issued debt. A key indicator in the model will be the debt service coverage ratio (DSCR). The more conservative the DSCR, the more comfort there will be for investors in the notes. For a residential mortgage deal, this ratio might be approximately 2.5–3.0; however for an airline ticket receivables deal, the DSCR would be unlikely to be lower than 4.0. The model will therefore calculate the amount of notes that can be issued against the assets, whilst maintaining the minimum DSCR.

Credit rating

It is common for securitisation deals to be rated by one or more of the formal credit ratings agencies such as Moody's, Fitch or Standard & Poor's. A formal credit rating will make it easier for XYZ Securities to place the notes with investors. The methodology employed by the ratings agencies takes into account both qualitative and quantitative factors, and will differ according to the asset class being securitised. The main issues in a deal such as our hypothetical Airway No 1 deal would be expected to include:

- *Corporate credit quality*. These are risks associated with the originator, and are factors that affect its ability to continue operations, meet its financial obligations, and provide a stable foundation for generating future receivables. This might be analysed according to the following: first, ABC Airways' historical financial performance, including its liquidity and debt structure; second, its status within its domicile country, for example whether it is state owned; third, the general economic conditions for industry and for airlines; and fourth, the historical record and current state of the airline, for instance its safety record and age of its airplanes;
- *The competition and industry trends*: ABC Airways' market share, the competition on its network.

- *Regulatory issues*, such as need for ABC Airways to comply with forthcoming legislation that would impact its cash flows.
- The *legal structure* of the SPV and transfer of assets.
- *Cash flow* analysis.

Based on the findings of the ratings agency, the arranger may redesign some aspect of the deal structure so that the issued notes are rated at the required level.

This is a selection of the key issues involved in the process of securitisation. Depending on investor sentiment, market conditions and legal issues, the process from inception to closure of the deal may take anything from three to 12 months or more. After the notes have been issued, the arranging bank will no longer have anything to do with the issue; however the bonds themselves require a number of agency services for their remaining life until they mature or are paid off. These agency services include paying agent, cash manager and custodian.

Securitisation post-credit crunch

Following the July–August 2007 implosion of the asset-backed commercial paper market, investor interest in ABS product dried up virtually completely. The growing illiquidity in the interbank market, which resulted in even large AA-rated banks finding it difficult to raise funds for tenors longer than one month, became acute following the collapse of Lehman Brothers in September 2008. To assist banks in raising funds, central banks starting with the US Federal Reserve and European Central Bank (ECB), and then the Bank of England (BoE), began to relax the criteria under which they accepted collateral from banks that raised term funds from them. In summary, the central banks announced that asset-backed securities, including mortgage-backed securities, would now be eligible as collateral at the daily liquidity window.

As originally conceived, the purpose of these moves was to enable banks to raise funds, from their respective central banks, using existing ABS on their balance sheets as collateral. Very quickly, however, the banks began to originate new securitisation transactions; for instance a bank would use illiquid assets held on its balance sheet (such as residential mortgages or corporate loans) as collateral in the deal. The issued notes would be purchased by the bank itself, making the deal completely in-house. These new purchased ABS tranches would then be used as collateral at the central bank repo window.

Structuring considerations

Essentially an ECB-led deal is like any other ABS deal except that there is a minimum requirement to be ECB eligible. There are also haircut considerations, and the opportunity to structure the deal without consideration for investors' requirements. To be eligible for repo at the ECB, deals had to fulfil certain criteria. These included:

(a) minimum requirements:
    - public rating of AAA or higher at first issue

- only the senior tranche can be repo'd
- No exposure to synthetic securities. The ECB rules state that the cash flow in generating assets backing the asset-backed securities must not consist in whole or in part, actually or potentially, of credit-linked notes or similar claims resulting from the transfer of credit risk by means of credit derivatives. Therefore, the transaction should expressly exclude any types of synthetic assets or securities
- public presale or new issue report issued by the rating agency that is rating the facility
- listed in Europe (for example, the Irish Stock Exchange)
- book entry capability in Europe (for example Euroclear, Clearstream).

(b) haircut considerations:

- Collateralised loan obligations (CLO) securities denominated in euro will (from March 2009), incur a haircut of 12% regardless of maturity or coupon structure.
- For the purposes of valuation, in the absence of a trading price within the past five days, or if the price is unchanged over that period, a 5% valuation markdown is applied. This equates to an additional haircut of 4.4%
- CLO securities denominated in USD will incur the usual haircuts but with an additional initial margin of between 10% and 20% to account for FX risk.

(c) other considerations:

- can incorporate a revolving period (external investors normally would not prefer this)
- can be a simple two-tranche set up. The junior tranche can be unrated and subordinated to topping up the cash reserve
- off-market swap (out of money on day one) allowed
- one rating agency (note that the BoE requires two)
- there can be no in-house currency swap (this must be with an external counterparty).

The originator also must decide whether the transaction is to be structured to accommodate replenishment of the portfolio or whether the portfolio should be static. ECB transactions are clearly financing transactions for the bank, and as such the bank will wish to retain flexibility to sell or refinance some or all of the portfolio at any time should more favourable financing terms become available to it. For this reason there is often no restriction on the ability to sell assets out of the portfolio provided that the price received by the issuer is not less than the price paid by it for the asset (par), subject to adjustment for accrued interest. This feature maintains maximum refinancing flexibility and has been agreed to by the rating agencies in previous transactions.

Whether or not replenishment is incorporated into the transaction depends on a number of factors. If it is considered likely that assets will be transferred out of the portfolio (in order to be sold or refinanced), then replenishment enables the efficiency of the ABS structure to be maintained by adding new assets rather than running the existing transaction down and having to establish a new structure to finance additional/future assets. However, if replenishment is incorporated into

the transaction, the rating agencies will have to carry out diligence on the bank to satisfy themselves of the capabilities of the bank to manage the portfolio. Also the recovery rates assigned to a static portfolio will be higher than those assigned to a manager portfolio. The decision on whether to have a managed or static transaction will impact on the documentation for the transaction, and the scope of the bank's obligations and representations.

Example of an in-house deal

During 2007–9 over 100 banks in the European Union undertook in-house securitisations in order to access the ECB discount window, as funding sources in the interbank market dried up.[6] A United Kingdom banking institution, the Nationwide Building Society, acquired an Irish banking entity during 2008 essentially in order to access the ECB's discount window (for which a branch or subsidiary in the euro-zone area was a requirement).

One such public deal was Fast Net Securities 3 Limited, originated by Irish Life and Permanent plc. Example 12.1 shows the deal highlights.

Example 12.1 Fast Net Securities 3 Limited

Fast Net Securities 3 Limited

Class	*Balance*	*% of total*	*Rating (S&P)*	*WAL (years)*	*Legal final*	*Basis*	*Margin (bp)*
A1	1,920,000,000	24%	AAA	2.91	Nov-2049	1-mo Euribor	40
A2	5,040,000,000	63%	AAA	3.15	Nov-2049	1-mo Euribor	45
B	1,040,000,000	13%	n/r	3.08	Nov-2049	1-mo Euribor	200
	8,000,000,000						
Cash reserve	400,000,000	5%					
Swap spread							150

[6] The entire business model of a large number of banks as well as 'shadow banks' such as Structured Investment Vehicles (SIVs) had depended on available liquidity from the interbank market, which was rolled over on a short-term basis such as weekly or monthly and used to fund long-dated assets such as RMBS securities that had much longer maturities and which themselves could not be realised in a liquid secondary market once the 2007 credit crunch took hold. This business model unravelled after the credit crunch, with its most notable casualties being Northern Rock plc and the SIVs themselves, which collapsed virtually overnight. Regulatory authorities responded by requiring banks to take liquidity risk more seriously, with emphasis on longer-term average tenor of liabilities and greater diversity of funding sources (for example see the UK FSA's CP 08/22 document, at www.fsa.org). The author discusses liquidity management in *Bank Asset and Liability Management* (Wiley Asia 2007) and *The Principles of Banking* (Wiley Asia 2010).

Timing

Cut-off date	12/5/07
Final offering circular	12/17/07
Settlement	12/17/07
First payment date	2/11/08

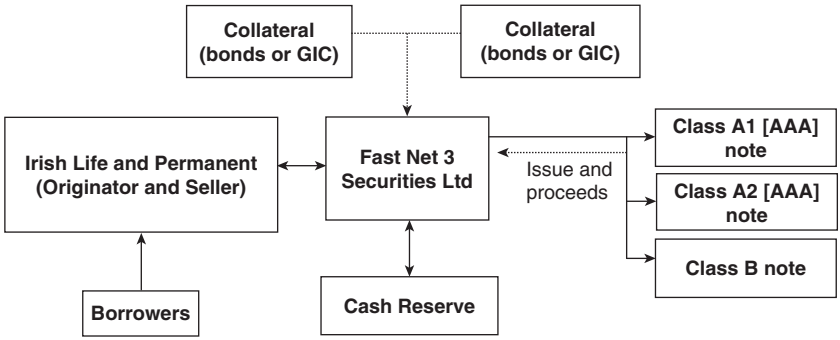
Key terms

Issuer	Fast Net Securities 3 Ltd
Originator	Irish Life and Permanent
Sole arranger	Deutsche Bank AG
Trustee	Deutsche Trustee Company Ltd
First payment date	February 11, 2008
Day count	Actual/360
Listing	Irish Stock Exchange
Settlement	Euroclear / Clearstream
Legal maturity date	November 11, 2049

Asset pool

Mortgage pool	Residential mortgages originated by Irish Life in the Republic of Ireland
Number of obligors	35,672
Aggregate balance	EUR 8,319,049,200.22
Average balance	EUR 226,190
Largest mortgage	EUR 8,502,202
Weighted average Loan-to-Value Ratio (LTV)	83%
Weighted average seasoning	21 months
Weighted average remaining term	27.3 years
Longest maturity date	2-Nov-47

Structure diagram



(*Source*: S&P. Details reproduced with permission).

Note that this transaction was closed in December 2007, a time when the securitisation market was essentially moribund in the wake of the credit crunch. An ABS note rated AAA could be expected to be marked-to-market at over 200 bps

over Libor. Because the issued notes were purchased in entirety by the originator, who intended to use the senior tranche as collateral to raise funds at the ECB, the terms of the deal could be set at a purely nominal level; this explains the '40 bps over Euribor' coupon of the senior tranche

References and bibliography

Bhattacharya, A. and Fabozzi, F. (eds) *Asset-Backed Securities*, FJF Associates, 1996.
CFSB, *Credit Risk Transfer*, 2 May 2003.
Choudhry, M. *The Bond and Money Markets: Strategy, Trading, Analysis*, Butterworth-Heinemann, 2001.
Choudhry, M. *Bank Asset and Liability Management*, Wiley Asia (2007).
Choudhry, M. *Structured Credit Products,* 2nd edition, Wiley Asia (2010).
Hayre, L. (ed.), *The Salomon Smith Barney Guide to Mortgage-Backed and Asset-Backed Securities*, Wiley, 2001.
Martellini, L., Priaulet, P. and Priaulet, S. *Fixed Income Securities*, Wiley, 2003.
Morris, D. *Asset Securitisation: Principles and Practices*, Executive Enterprise, 1990.
Sundaresan, S. *Fixed Income Markets and their Derivatives*, South-Western Publishing, 1997, ch. 9.

13
Mortgage-Backed Securities

In Chapter 12 we introduced *asset-backed bonds*, debt instruments created from a package of loan assets on which interest is payable, usually on a floating basis. The asset-backed market was developed in the United States and is a large, diverse market containing a wide variety of instruments. The characteristics of asset-backed securities (ABS) present additional features in their analysis, which are investigated in this and the next two chapters. Financial engineering techniques employed by investment banks today enable an entity to create a bond structure from any type of cash flow; the typical forms are high volume loans such as residential mortgages, car loans and credit card loans. The loans form assets on a bank or finance house balance sheet, which are packaged together and used as backing for an issue of bonds. The interest payments on the original loans form the cash flows used to service the new bond issue.

In this chapter we consider *mortgage-backed securities,* part of the asset-backed securities market.

A *mortgage* is a loan made for the purpose of purchasing property, which in turn is used as the security for the loan itself. It is defined as a debt instrument giving conditional ownership of an asset, and secured by the asset that is being financed. The borrower provides the lender with a mortgage in exchange for the right to use the property during the term of the mortgage, and agrees to make regular payments of both principal and interest. The mortgage lien is the security for the lender, and is removed when the debt is paid off. A mortgage may involve residential property or commercial property and is a long-term debt, normally 25 to 30 years; however it can be drawn up for shorter periods if required by the borrower. If the borrower or *mortgagor* defaults on the interest payments, the lender or *mortgagee* has the right to take over the property and recover the debt from the proceeds of selling the property. Mortgages can carry either *fixed-rate* or *floating-rate* interest. Although in the United States mortgages are generally *amortising loans*, known as *repayment mortgages* in the UK, there are also *interest-only mortgages* where the borrower only pays the interest on the loan. On maturity the original loan amount is paid off by the proceeds of a maturing investment contract taken out at the same time as the mortgage. These are known as *endowment mortgages* and are popular in the UK market, although their popularity has been waning in recent years.

A lending institution may have many hundreds of thousands of individual residential and commercial mortgages on its book. If the total loan book is pooled together and used as collateral for the issue of a bond, in a securitisation process, the resulting instrument is a *mortgage-backed security*.

In the US market certain mortgage-backed securities are backed, either implicitly or explicitly, by the government, in which case they trade essentially as risk-free instruments and are not rated by the credit agencies. In the United States government agencies, the Government National Mortgage Association (GNMA, known as 'Ginnie Mae'), the Federal Home Loan Corporation and the Federal National Mortgage Association ('Freddie Mac' and 'Fannie Mae' respectively), purchase mortgages for the purpose of pooling them and holding them in their portfolios; they may then be securitised. Bonds that are not issued by government agencies are rated in the same way as other corporate bonds. On the other hand non-government agencies sometimes obtain mortgage insurance for their issue, in order to boost its credit quality. When this happens the credit rating of the mortgage insurer becomes an important factor in the credit standing of the bond issue.

Mortgages

In the US market, the terms of a conventional mortgage, known as a *level-payment fixed-rate mortgage*, will state the interest rate payable on the loan, the term of the loan and the frequency of payment. Most mortgages specify monthly payment of interest. These are in fact the characteristics of a level-payment mortgage, which has a fixed interest rate and fixed term to maturity. This means that the monthly interest payments are fixed, hence the term 'level-pay'.

The singular feature of a mortgage is that, even if it charges interest at a fixed rate, its cash flows are not known with absolute certainty. This is because the borrower can elect to repay any or all of the principal before the final maturity date. This is a characteristic of all mortgages, and although some lending institutions impose a penalty on borrowers who retire the loan early, this is an investment risk for the lender, known as *prepayment risk*.

The monthly interest payment on a conventional fixed-rate mortgage is given by (13.3), which is derived from the conventional present value analysis used for an annuity. Essentially the primary relationship is:

$$M_{m0} = I\left(\frac{1-1/(1+r)^n}{r}\right) \tag{13.1}$$

from which we can derive:

$$I = \frac{M_{m0}}{\dfrac{1-1/(1+r)^n}{r}} \tag{13.2}$$

This is simplified to:

$$I = M_{m0} \frac{r(1+r)^n}{(1+r)^n - 1} \tag{13.3}$$

where

M_{m0} is the original mortgage balance (the cash amount of loan)
I is the monthly cash mortgage payment
r is the simple monthly interest rate, given by (annual interest rate/12)
n is the term of the mortgage in months.

The monthly repayment includes both the interest servicing and a repayment of part of the principal. In Example 13.1 after the 264th interest payment, the balance will be zero and the mortgage will have been paid off. Since a portion of the original balance is paid off every month, the interest payment reduces by a small amount each month, that is, the proportion of the monthly payment dedicated to repaying the principal steadily increases. The remaining mortgage balance for any particular month during the term of the mortgage may be calculated using (13.4):

$$M_{mt} = M_{m0} \frac{r(1+r)^n - r(1+r)^t}{(1+r)^n - 1} \tag{13.4}$$

where M_{mt} is the mortgage cash balance after t months and n remains the original maturity of the mortgage in months.

The level of interest payment and principal repayment in any one month during the mortgage term can be calculated using the equations below. If we wish to calculate the value of the principal repayment in a particular month during the mortgage term, we may use (13.5):

$$P_t = M_{m0} \frac{(1+r)^{t-1}}{(1+r)^n - 1} \tag{13.5}$$

where P_t is the scheduled principal repayment amount for month t, while the level of interest payment in any month is given by (13.6):

$$i_t = M_{m0} \frac{r((1+r)^n - (1+r)^{t-1})}{(1+r)^n - 1} \tag{13.6}$$

where i_t is the interest payment only in month t.

Some mortgage contracts incorporate a *servicing fee*. This is payable to the mortgage provider to cover the administrative costs associated with collecting interest payments, sending regular statements and other information to borrowers, chasing

Example 13.1: Mortgage contract calculations

A mortgage borrower enters into a conventional mortgage contract, in which he borrows £72,200 for 22 years at a rate of 7.99%. What is the monthly mortgage payment?

This gives us n equal to 264 and r equal to (0.0799/12) or 0.0066583. Inserting the above terms into (13.3) we have:

$$I = 72{,}200\left\{\frac{0.0066583(1.0066583)^{264}}{(1.0066583)^{264}-1}\right\}$$

or I equal to £581.60

The mortgage balance after ten years is given below, where t is 120:

$$M_{m120} = \left\{\frac{(1.0066583)^{264}-(1.0066583)^{120}}{(1.0066583)^{264}-1}\right\}$$

or a remaining balance of £53,756.93.

In the same month the scheduled principal repayment amount is:

$$P_{120} = 72{,}200\left\{\frac{0.0066583(1.0066583)^{120-1}}{(1.0066583)^{264}-1}\right\}$$

or £222.19.

The interest only payable in month 120 is shown below:

$$i_{120} = 72{,}200\left\{\frac{0.0066583(1.0066583)^{264}-(1.0066583)^{120-1}}{(1.0066583)^{264}-1}\right\}$$

and is equal to £359.41. The combined mortgage payment is £581.60, as calculated before.

overdue payments, maintaining the records and processing systems and other activities. Mortgage providers also incur costs when repossessing properties after mortgagors have fallen into default. Mortgages may be serviced by the original lender or another third-party institution that has acquired the right to service them, in return for collecting the fee. When a servicing charge is payable by a borrower, the monthly mortgage payment is comprised of the interest costs, the principal repayment and the servicing fee. The fee incorporated into the monthly payment is usually stated as a percentage, say 0.25%. This is added to the mortgage rate.

Another type of mortgage in the US market is the *adjustable-rate mortgage* or ARM, which is a loan in which the interest rate payable is set in line with an external reference rate. The re-sets are at periodic intervals depending on the terms of

the loan, and can be on a monthly, six-monthly or annual basis, or even longer. The interest rate is usually fixed at a spread over the reference rate. The reference rate that is used can be a market-determined rate such as the prime rate, or a calculated rate based on the funding costs for US savings and loan institutions or thrifts. The cost of funds for thrifts is calculated using the monthly average funding cost on the thrifts' activities, and there are 'thrift indexes' that are used to indicate the cost of funding. The two most common indices are the Eleventh Federal Home Loan Bank Board District Cost of Funds Index (COFI) and the National Cost of Funds Index. Generally borrowers prefer to fix the rate they pay on their loans to reduce uncertainty, and this makes fixed-rate mortgages more popular than variable rate mortgages. A common incentive used to entice borrowers away from fixed-rate mortgages is to offer a below-market interest rate on an ARM mortgage, usually for an introductory period. This comfort period may be from two to five years or even longer.

Mortgages in the UK are predominantly *variable-rate mortgages*, in which the interest rate moves in line with the clearing bank base rate. It is rare to observe fixed-rate mortgages in the UK market, although short-term fixed-rate mortgages are more common (the rate reverts to a variable basis at the termination of the fixed-rate period).

A *balloon mortgage* entitles a borrower to long-term funding, but under its terms, at a specified future date the interest rate payable is renegotiated. This effectively transforms a long-dated loan into a short-term borrowing. The balloon payment is the original amount of the loan, minus the amount that is amortised. In a balloon mortgage therefore the actual maturity of the bonds is below the stated maturity.

A *graduated payment mortgage* (GPM) is aimed at lower-earning borrowers, as the mortgage payments for a fixed initial period, say the first five years, are set at lower than the level applicable for a level-paying mortgage with an identical interest rate. The later mortgage payments are higher as a result. Hence a GPM mortgage will have a fixed term and a mortgage rate, but the offer letter will also contain details on the number of years over which the monthly mortgage payments will increase and the point at which level payments will take over. There will also be information on the annual increase in the mortgage payments. As the initial payments in a GPM are below the market rate, there will be little or no repayment of principal at this time. This means that the outstanding balance may actually increase during the early stages, a process known as *negative amortisation*. The higher payments in the remainder of the mortgage term are designed to pay off the entire balance in maturity. The opposite to the GPM is the *growing equity mortgage* or GEM. This mortgage charges fixed-rate interest but the payments increase over time; this means that a greater proportion of the principal is paid off over time, so that the mortgage itself is repaid in a shorter time than the level-pay mortgage.

In the UK market it is more common to encounter hybrid mortgages, which charge a combination of fixed-rate and variable-rate interest. For example the rate may be fixed for the first five years, after which it will vary with changes in the lender's base rate. Such a mortgage is known as a *fixed/adjustable hybrid mortgage*.

Mortgage risk

Although mortgage contracts are typically long-term loan contracts, running usually for 20 to 30 years or even longer, there is no limitation on the amount of the principal that may be repaid at any one time. In the US market there is no penalty for repaying the mortgage ahead of its term, known as a *mortgage prepayment*. In the UK some lenders impose a penalty if a mortgage is prepaid early, although this is more common for contracts that have been offered at special terms, such as a discounted loan rate for the start of the mortgage's life. The penalty is often set as extra interest, for example six months' worth of mortgage payments at the time when the contract is paid off. As borrowers are free to prepay a mortgage at a time of their choosing, the lender accepts a prepayment risk, as noted earlier.

A borrower may pay off the principal ahead of the final termination date for a number of reasons. The most common reason is when the property on which the mortgage is secured is subsequently sold by the borrower; this results in the entire mortgage being paid off at once. The average life of a mortgage in the UK market is eight years, and mortgages are most frequently prepaid because the property has been sold (source: *HBOS plc*). Other actions that result in the prepayment of a mortgage are when a property is repossessed after the borrower has fallen into default, if there is a change in interest rates making it attractive to refinance the mortgage (usually with another lender), or if the property is destroyed due to accident or natural disaster.

An investor acquiring a pool of mortgages from a lender will be concerned at the level of prepayment risk, which is usually measured by projecting the level of expected future payments using a financial model. Although it would not be possible to evaluate meaningfully the potential of an individual mortgage to be paid off early, it is tenable to conduct such analysis for a large number of loans pooled together. A similar activity is performed by actuaries when they assess the future liability of an insurance provider who has written personal pension contracts. Essentially the level of prepayment risk for a pool of loans is lower than that of an individual mortgage. Prepayment risk has the same type of impact on a mortgage pool's performance and valuation as a call feature does on a callable bond. This is understandable because a mortgage is essentially a callable contract, with the 'call' at the option of the borrower of funds.

The other significant risk of a mortgage book is the risk that the borrower will fall into arrears, or be unable to repay the loan on maturity. This is known as *default risk*. Lenders take steps to minimise the level of default risk by assessing the credit quality of each borrower, as well as the quality of the property itself. A study (Brown *et al.*, 1990) has found that the higher the deposit paid by the borrower, the lower the level of default. Therefore lenders prefer to advance funds against a borrower's equity that is deemed sufficient to protect against falls in the value of the property. In the UK the typical deposit required is 25%, although certain lenders will advance funds against smaller deposits such as 10% or 5%.

Securities

Mortgage-backed securities are bonds created from a pool of mortgages. They are formed from mortgages that are for residential or commercial property or a mixture of both. Bonds created from commercial mortgages are known as *commercial mortgage-backed securities*. There are a range of different securities in the market, known in the United States as *mortgage pass-through securities*. There also exist two related securities known as *collateralised mortgage securities* and *stripped mortgage-backed securities*. Bonds that are created from 'mortgage pools' that have been purchased by government agencies are known as *agency mortgage-backed securities*, and are regarded as risk-free in the same way as Treasury securities.

A mortgage-backed bond is created by an entity out of its mortgage book or a book that it has purchased from the original lender (there is very often no connection between a mortgage-backed security and the firm that made the original loans). The mortgage book will have a total nominal value comprised of the total value of all the individual loans. The loans will generate cash flows, consisting of the interest and principal payments, and any prepayments. The regular cash flows are received on the same day each month, so the pool resembles a bond instrument. Therefore bonds may be issued against the mortgage pool. Example 13.2 is a simple illustration of a type of mortgage-backed bond known as a mortgage pass-through security in the US market.

Example 13.2: Mortgage pass-through security

An investor purchases a book consisting of 5000 individual mortgages, with a total repayable value of US$500,000,000. The loans are used as collateral against the issue of a new bond, and the cash flows payable on the bond are the cash flows that are received from the mortgages. The issuer sells 1000 bonds, with a face value of US$500,000. Each bond is therefore entitled to 1/1000 or 0.02% of the cash flows received from the mortgages.

The prepayment risk associated with the original mortgages is unchanged, but any investor can now purchase a bond with a much lower value than the mortgage pool but with the same level of prepayment risk, which is lower than the risk of an individual loan. This would have been possible if an investor was buying all 100 mortgages, but by buying a bond that represents the pool of mortgages, a smaller cash value is needed to achieve the same performance. The bonds will also be more liquid than the loans, and the investor will be able to realise her investment ahead of the maturity date if she wishes. For these reasons the bonds will trade at higher prices than would an individual loan. A mortgage pass-through security therefore is a way for mortgage lenders to realise additional value from their loan book, and if it is sold to another investor (who issues the bonds), the loans will be taken off the original lender's balance sheet, thus freeing up lending lines for other activities.

A *collateralised mortgage obligation* (CMO) differs from a pass-through security in that the cash flows from the mortgage pool are distributed on a prioritised basis, based on the class of security held by the investor. In Example 13.2 this might mean that three different securities are formed, with a total nominal value of US$100 million each, entitled to a pro-rata amount of the interest payments but with different priorities for the repayment of principal. For instance, US$60 million of the issue might consist of a bond known as 'class A' which is entitled to receipt of all the principal repayment cash flows, after which the next class of bonds is entitled to all the repayment cash flow; these would be 'class B' bonds, of which say $25 million worth were created, and so on. If 300 class A bonds are created, they would have a nominal value of $200,000 and each would receive 0.33% of the total cash flows received by the class A bonds. Note that all classes of bonds receive an equal share of the interest payments; it is the principal repayment cash flows received that differ. What is the main effect of this security structure? The most significant factor is that, in our illustration, the class A bonds will be repaid earlier than any other class of bond that is formed from the securitisation. They therefore have the shortest maturity. The last class of bonds will have the longest maturity. There is still a level of uncertainty associated with the maturity of each bond, but this is less than the uncertainty associated with a pass-through security.

Cash flow patterns

We stated that the exact term of a mortgage-backed bond cannot be stated with accuracy at the time of issue, because of the uncertain frequency of mortgage prepayments. This uncertainty means that it is not possible to analyse the bonds using the conventional methods used for fixed coupon bonds. The most common approach used by the market is to assume a fixed prepayment rate at the time of issue and use this to project the cash flows, and hence the lifespan, of the bond. The choice of prepayment selected is therefore significant, although it is also recognised that prepayment rates are not stable and will fluctuate with changes in mortgage rates and the economic cycle. In this section we consider some of the approaches used in evaluating the prepayment pattern of a mortgage-backed bond.

Prepayment analysis

Some market analysts assume a fixed life for a mortgage pass-through bond based on the average life of a mortgage. Traditionally a '12-year prepaid life' has been used to evaluate the securities, as market data suggested that the average mortgage has been paid off after the twelfth year. This is not generally favoured because it does not take into account the effect of mortgage rates and other factors. A more common approach is to use a constant prepayment rate (CPR). This measure is based on the expected number of mortgages in a pool that will be prepaid in a selected period, and is an annualised figure. The measure for the monthly level of prepayment is known as the *constant monthly repayment*, and measures the expected amount of the outstanding balance, minus the scheduled principal, that will be

prepaid in each month. Another name for the constant monthly repayment is the *single monthly mortality rate* or SMM. In Fabozzi (1997) the SMM is given by (13.7) and is an expected value for the percentage of the remaining mortgage balance that will be prepaid in that month.

$$SMM = 1 - (1 - CPR)^{1/12} \tag{13.7}$$

Example 13.3: Constant prepayment rate

The constant prepayment rate for a pool of mortgages is 2% each month. The outstanding principal balance at the start of the month is £72,200, while the scheduled principal payment is £223. This means that 2% of £71,977, or £1,439 will be prepaid in that month. To approximate the amount of principal prepayment, the constant monthly prepayment is multiplied by the outstanding balance.

In the US market the convention is to use the prepayment standard developed by the Public Securities Association (PSA), which is the domestic bond market trade association (since renamed the Bond Market Association and now known as the SIFMA). The PSA benchmark, known as 100% PSA, assumes a steadily increasing constant prepayment rate each month until the thirtieth month, when a constant rate of 6% is assumed. The starting prepayment rate is 0.2%, increasing by 0.2% each month until the rate levels off at 6%.

For the 100% PSA benchmark we may set, if t is the number of months from the start of the mortgage, that if $t < 30$, the CPR = 6%.t/30 while if $t > 30$, then the CPR is equal to 6%.

This benchmark can be altered if required to suit changing market conditions, so for example the 200% PSA has a starting prepayment rate and an increase that is double the 100% PSA model, so the initial rate is 0.4%, increasing by 0.4% each month until it reaches 12% in the thirtieth month, at which point the rate remains constant. The 50% PSA has a starting (and increases by a) rate of 0.1%, remaining constant after it reaches 3%.

The prepayment level of a mortgage pool will have an impact on its cash flows. As we saw in Example 13.1, if the amount of prepayment is nil, the cash flows will remain constant during the life of the mortgage. In a fixed-rate mortgage the proportion of principal and interest payment will change each month as more and more of the mortgage amortises. That is, as the principal amount falls each month, the amount of interest decreases. If we assume that a pass-through security has been issued today, so that its coupon reflects the current market level, the payment pattern will resemble the bar chart shown at Figure 13.1.

When there is an element of prepayment in a mortgage pool, for example as in the 100% PSA or 200% PSA model, the amount of principal payment will increase

during the early years of the mortgages and then becomes more steady, before declining for the remainder of the term. This is because the principal balance has declined to such an extent that the scheduled principal payments become less significant. Two examples are shown as Figures 13.2 and 13.3.

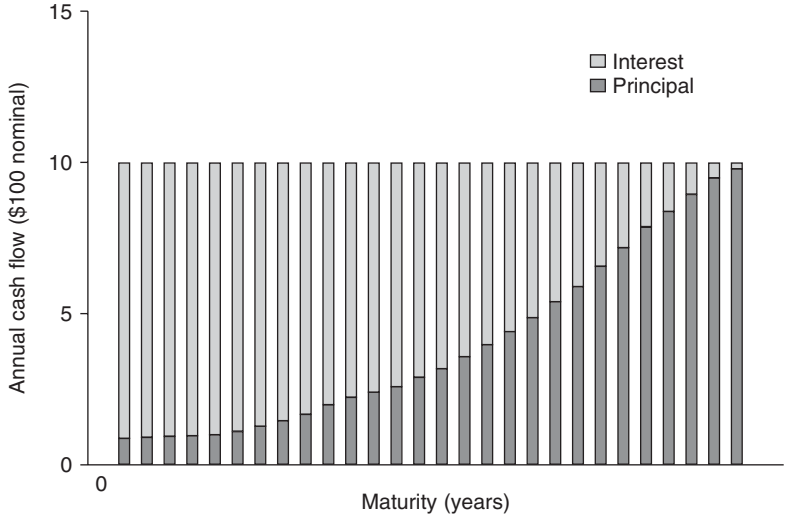


Figure 13.1 Mortgage pass-through security with 0% constant prepayment rate

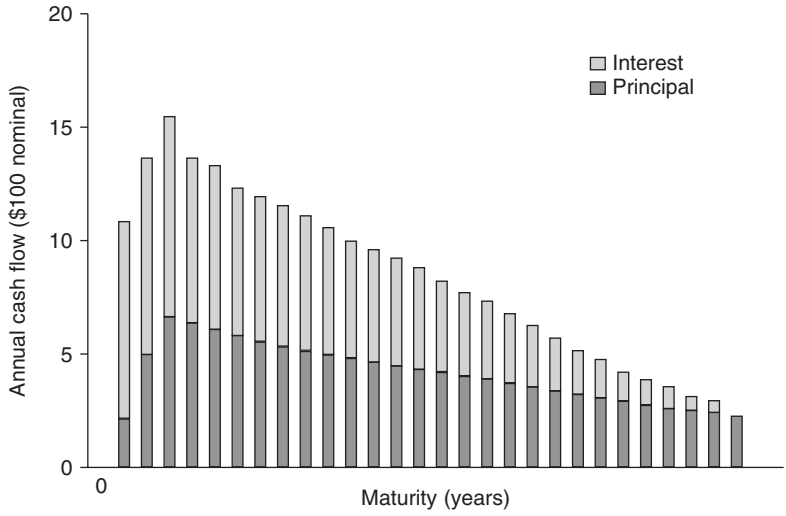


Figure 13.2 100% PSA model

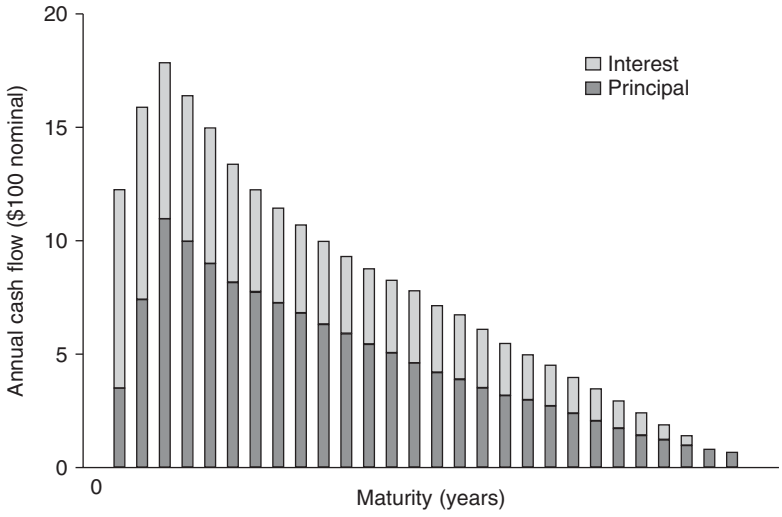


Figure 13.3 200% PSA model

The prepayment volatility of a mortgage-backed bond will vary according to the interest rate of the underlying mortgages. It has been observed that where the mortgages have interest rates of between 100 and 300 basis points above current mortgage rates, the prepayment volatility is the highest. At the bottom of the range, any fall in interest rates often leads to a sudden increase in refinancing of mortgages, while at the top of the range, an increase in rates will lead to a decrease in the prepayment rate.

The actual cash flow of a mortgage pass-through is of course dependent on the cash flow patterns of the mortgages in the pool. The relationships described in Example 13.1 can be used to derive further expressions to construct a cash flow schedule for a pass-through security, using a constant or adjustable assumed prepayment rate. Fabozzi (1997) describes the projected monthly mortgage payment for a level-paying fixed rate mortgage in any month as

$$\bar{I}_t = \overline{M}_{mt-1} \frac{r(1+r)^{n-t+1}}{(1+r)^{n-t+1}-1} \tag{13.8}$$

where

$\bar{I}_t$ is the projected monthly mortgage payment for month t

$\overline{M}_{mt-1}$ is the projected mortgage balance at the end of month t assuming that prepayments have occurred in the past.

To calculate the interest proportion of the projected monthly mortgage payment we use (13.9) where it is the projected monthly interest payment for month t.

$$\bar{I}_t = \overline{M}_{mt-1}.i. \tag{13.9}$$

Expression (13.9) states that the projected monthly interest payment can be obtained by multiplying the mortgage balance at the end of the previous month by the monthly interest rate. In the same way the expression for calculating the projected monthly scheduled principal payment for any month is given by (13.10), where $\overline{P}_t$ is the projected scheduled principal payment for the month *t*.

$$\overline{P}_t = \overline{I}_t - \overline{i}_t \tag{13.10}$$

The projected monthly principal prepayment, which is an expected rate only and not a model forecast, is given by (13.11):

$$\overline{P}\,\overline{P}_t = SMM_t(\overline{M}_{mt-1} - \overline{P}_t) \tag{13.11}$$

where $\overline{P}\overline{P}_t$ is the projected monthly principal prepayment for month *t*.

The above relationships enable us to calculate values for:

- the projected monthly interest payment
- the projected monthly scheduled principal payment
- the projected monthly principal prepayment.

These values may be used to calculate the total cash flow in any month that a holder of a mortgage-backed bond receives, which is given by (13.12), where cf_t is the cash flow receipt in month *t*.

$$cf_t = \overline{I}_t + \overline{P}_t + \overline{P}\overline{P}_t \tag{13.12}$$

The practice of using a prepayment rate is a market convention that enables analysts to evaluate mortgage-backed bonds. The original PSA prepayment rates were arbitrarily selected, based on the observation that prepayment rates tended to stabilise after the first 30 months of the life of a mortgage. A linear increase in the prepayment rate is also assumed. However this is a market convention only, adopted by the market as a standard benchmark. The levels do not reflect seasonal variations in prepayment patterns, or the different behaviour patterns of different types of mortgages.

The PSA benchmarks can be (and are) applied to default assumptions to produce a default benchmark. This is used for non-agency mortgage-backed bonds only, as agency securities are guaranteed by one of the three government or government-sponsored agencies. Accordingly the PSA standard default assumption (SDA) benchmark is used to assess the potential default rate for a mortgage pool. For example the standard benchmark, 100SDA assumes that the default rate in the first month is 0.02% and increases in a linear fashion by 0.02% each month until the thirtieth month, at which point the default rate remains at 0.60%. In month 60 the default rate begins to fall from 0.60% to 0.03% and continues to fall linearly until month 120. From that point the default rate remains constant at 0.03%. The other benchmarks have similar patterns.

Prepayment models

The PSA standard benchmark reviewed in the previous section uses an assumption of prepayment rates, and can be used to calculate the prepayment proceeds of a mortgage. It is not, strictly speaking, a prepayment model because it cannot be used to estimate actual prepayments. A prepayment model on the other hand does attempt to predict the prepayment cash flows of a mortgage pool, by modelling the statistical relationships between the various factors that have an impact on the level of prepayment. These factors are the current mortgage rate, the characteristics of the mortgages in the pool, seasonal factors and the general business cycle. Let us consider them in turn.

The prevailing mortgage interest rate is probably the most important factor in the level of prepayment. The level of the current mortgage rate and its spread above or below the original contract rate will influence the decision to refinance a mortgage; if the rate is materially below the original rate, the borrower will prepay the mortgage. As the mortgage rate at any time reflects the general bank base rate, the level of market interest rates has the greatest effect on mortgage prepayment levels. The current mortgage rate also has an effect on housing prices, since if mortgages are seen as 'cheap' the general perception will be that now is the right time to purchase: this affects housing market turnover. The pattern followed by mortgage rates since the original loan also has an impact, a phenomenon known as *refinancing burnout*.

Evaluation and analysis of mortgage-backed bonds

Term to maturity

The term to maturity cannot be given for certain for a mortgage pass-through security, since the cash flows and prepayment patterns cannot be predicted. To evaluate such a bond therefore it is necessary to estimate the term for the bond, and use this measure for any analysis. The maturity measure for any bond is important, as without this it is not possible to assess over what period of time a return is being generated; also, it will not be possible to compare the asset with any other bond. The term to maturity of a bond also gives an indication of its sensitivity to changes in market interest rates. If comparisons with other securities such as government bonds are made, we cannot use the stated maturity of the mortgage-backed bond because prepayments will reduce this figure. The convention in the market is to use other estimated values, which are *average life* and the more traditional duration measure.

The average life of a mortgage pass-through security is the weighted average time to return of a unit of principal payment, made up of projected scheduled principal payments and principal prepayments. It is also known as the *weighted average life*. It is given by (13.13):

$$\text{Average life} = \frac{1}{12}\sum_{t=1}^{n}\frac{t(\text{Principal received at } t)}{\text{Total principal received}} \tag{13.13}$$

where n is the number of months remaining. The time from the term measured by the average life to the final scheduled principal payment is the bond's *tail*.

To calculate duration (or Macaulay's duration) for a bond we require the weighted present values of all its cash flows. To apply this for a mortgage-backed bond therefore it is necessary to project the bond's cash flows, using an assumed prepayment rate. The projected cash flows, together with the bond price and the periodic interest rate, may then be used to arrive at a duration value. The periodic interest rate is derived from the yield. This calculation for a mortgage-backed bond produces a periodic duration figure, which must be divided by 12 to arrive at a duration value in years (or by 4 in the case of a quarterly-paying bond).

Example 13.4: Macaulay duration

A 25-year mortgage security with a mortgage rate of 8.49% and monthly coupon is quoted at a price of US$98.50, a bond-equivalent yield of 9.127%. To calculate the Macaulay duration we require the present value of the expected cash flows using the interest rate that will make this present value, assuming a constant prepayment rate, equate the price of 98.50. Using the expression below:

$$rm = 2((1 + r)^n - 1)$$

where rm is 9.127% and $n = 5$, this is shown to be 9.018%.

For the bond above this present value is 6,120.79. Therefore the mortgage security Macaulay duration is given by:

$$D_m = \frac{6{,}120.79}{98.50} = 6214$$

Therefore the bond-equivalent Macaulay duration in years is given by

$$D = \frac{6214}{12} = 5.178$$

Calculating yield and price: static cash flow model

There are a number of ways that the yield on a mortgage-backed bond can be calculated. One of the most common methods employs the *static cash flow model*. This assumes a single prepayment rate to estimate the cash flows for the bond, and does not take into account how changes in market conditions might impact the prepayment pattern.

The conventional yield measure for a bond is the discount rate at which the sum of the present values of all the bond's expected cash flows will be equal to the price of the bond. The convention is usually to compute the yield from the *clean price*, that is, excluding any accrued interest. This yield measure is known as the bond's *redemption yield* or *yield-to-maturity*. However for mortgage-backed bonds it is known as a *cash flow yield* or *mortgage yield*. The cash flow for a mortgage-backed bond is not known with certainty, due to the effect of prepayments, and so must be derived using an assumed prepayment rate. Once the projected cash flows have been calculated, it is possible to calculate the cash flow yield. The formula is given by (13.14):

$$P = \sum_{n=1}^{N} \frac{C(t)}{(1+ri/1200)^{t-1}} \tag{13.14}$$

Note however that a yield so computed will be for a bond with monthly coupon payments,[1] so it is necessary to convert the yield to an annualised equivalent before any comparisons are made with conventional bond yields. In the US and UK markets, the bond-equivalent yield is calculated for mortgage-backed bonds and measured against the relevant government bond yield, which (in both cases) is a semi-annual yield. Although it is reasonably accurate to simply double the yield of a semi-annual coupon bond to arrive at the annualised equivalent,[2] to obtain the bond equivalent yield for a monthly paying mortgage-backed bond we use (13.15):

$$rm = 2((1 + ri_M)^6 - 1) \tag{13.15}$$

where rm is the bond equivalent yield and ri_M is the interest rate that will equate the present value of the projected monthly cash flows for the mortgage-backed bond to its current price. The equivalent semi-annual yield is given by (13.16):

$$rm_{s/a} = (1 + ri_M)^6 - 1 \tag{13.16}$$

The cash flow yield calculated for a mortgage-backed bond in this way is essentially the redemption yield, using an assumption to derive the cash flows. As such the measure suffers from the same drawbacks as it does when used to measure the return of a plain vanilla bond, which are that the calculation assumes a uniform reinvestment rate for all the bond's cash flows and that the bond will be held to maturity. The same weakness will apply to the cash flow yield measure for a mortgage-backed bond. In fact the potential inaccuracy of the redemption yield measure is even greater with a mortgage-backed bond because the frequency of

[1] The majority of mortgage-backed bonds pay interest on a monthly basis, since individual mortgages usually do as well; certain mortgage-backed bonds pay on a quarterly basis.
[2] See Chapter 1 for the formulae used to convert yields from one convention basis to another.

interest payments is higher, which makes the reinvestment risk greater. The final yield that is returned by a mortgage-backed bond will depend on the performance of the mortgages in the pool, specifically the prepayment pattern.

Given the nature of a mortgage-backed bond's cash flows, the exact yield cannot be calculated. However it is common for market practitioners to use the cash flow yield measure and compare this with the redemption yield of the equivalent government bond. The usual convention is to quote the spread over the government bond as the main measure of value. When measuring the spread, the mortgage-backed bond is compared with the government security that has a similar duration, or a term to maturity similar to its average life.

It is possible to calculate the price of a mortgage-backed bond once its yield is known (or vice versa). As with a plain vanilla bond, the price is the sum of the present values of all the projected cash flows. It is necessary to convert the bond-equivalent yield to a monthly yield, which is then used to calculate the present value of each cash flow.

Bond price and option-adjusted spread

The behaviour of mortgage securities often resembles that of callable bonds, because effectively there is a call feature attached to them, in the shape of the prepayment option of the underlying mortgage holders. This option feature is the principal reason why it is necessary to use average life as the term to maturity for a mortgage security. It is frequently the case that the optionality of a mortgage-backed bond, and the volatility of its yield, have a negative impact on the bond holders. This is for two reasons: the actual yield realised during the holding period has a high probability of being lower than the anticipated yield, which was calculated on the basis of an assumed prepayment level, and mortgages are frequently prepaid at the time when the bondholder will suffer the most: that is, prepayments occur most often when rates have fallen, leaving the bondholder to reinvest repaid principal at a lower market interest rate.

These features combined represent the biggest risk to an investor of holding a mortgage security, and market analysts attempt to measure and quantify this risk. This is usually done using a form of OAS analysis. Under this approach the value of the mortgagor's prepayment option is calculated in terms of a basis point penalty that must be subtracted from the expected yield spread on the bond. This basis point value is calculated using a binomial model or a simulation model to generate a range of future interest rate paths, only some of which will cause a mortgagor to prepay her mortgage. The interest rate paths that would result in a prepayment are evaluated for their impact on the mortgage bond's expected yield spread over a government bond.[3] As OAS analysis takes account of the option

[3]The yield spread from OAS analysis is based on the discounted value of the expected cash flow using the government bond-derived forward rate. The yield spread of the cash flow yield to the government bond is based on yields-to-maturity. For this reason, the two spreads are not strictly comparable. The OAS spread is added to the entire yield curve, whereas a yield spread is a spread over a single point on the government bond yield curve.

feature of a mortgage-backed bond, it will be less affected by a yield change than the bond's yield spread. Assuming a flat yield curve environment, the relationship between the OAS and the yield spread is given by:

$$\text{OAS} = \text{Yield spread} - \text{Cost of option feature}$$

This relationship can be observed occasionally when yield spreads on current coupon mortgages widen during upward moves in the market. As interest rates fall, the cost of the option feature on a current coupon mortgage will rise, as the possibility of prepayment increases. Put another way, the option feature begins to approach being in-the-money. To adjust for the increased value of the option, traders will price in higher spreads on the bond, which will result in the OAS remaining more or less unchanged.

Effective duration and convexity

The modified duration of a bond measures its price sensitivity to a change in yield; the calculation is effectively a snapshot of one point in time. It also assumes that there is no change in expected cash flows as a result of the change in market interest rates. Therefore it is an inappropriate interest rate risk for a mortgage-backed bond, whose cash flows would be expected to change after a change in rates, due to the prepayment effect. Hence mortgage-backed bonds react differently to interest rate changes than do conventional bonds, because when rates fall, the level of prepayments is expected to rise (and vice versa). Therefore when interest rates fall, the duration of the bond is likely also to fall, which is opposite to the behaviour of a conventional bond. This feature is known as *negative convexity* and is similar to the effect displayed by a callable bond. The prices of both these types of security react to interest rate changes differently than the price of conventional bonds.

For this reason the more accurate measure of interest rate sensitivity to use is effective duration as described by Fabozzi (1997). Effective duration is the approximate duration of a bond as given by (13.17):

$$D_{app} = \frac{P_- - P_+}{2P_0(\Delta rm)} \tag{13.17}$$

where

P_0 is the initial price of the bond
P_- is the estimated price of the bond if the yield decreases by Δrm
P_+ is the estimated price of the bond if the yield increases by Δrm
Δrm is the change in the yield of the bond.

The approximate duration is the effective duration of a bond when the two values P_- and P_+ are obtained from a valuation model that incorporates the effect of a change in the expected cash flows (from prepayment effects) when there is a change in interest rates. The values are obtained from a pricing model such as the

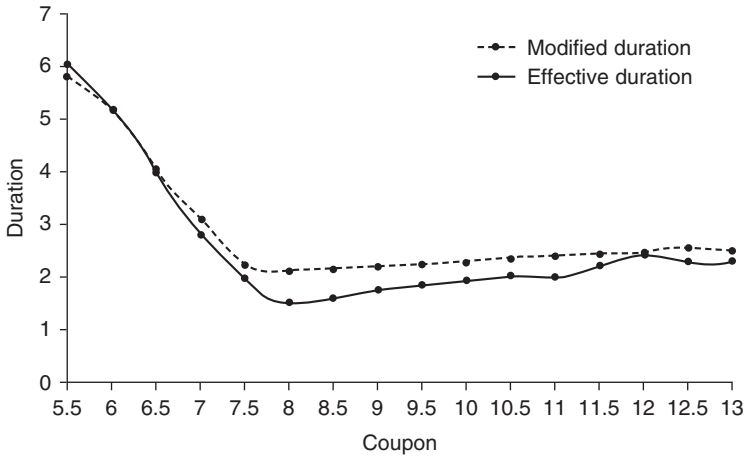


Figure 13.4 Modified duration and effective duration for agency mortgage-backed bonds

static cash flow model, binomial model or simulation model. The calculation of effective duration uses higher and lower prices that are dependent on the prepayment rate that is assumed. Generally analysts will assume a higher prepayment rate when the interest rate is at the lower level of the two.

Figure 13.4 illustrates the difference between modified duration and effective duration for a range of agency mortgage pass-through securities, where the effective duration for each bond is calculated using a 20 basis point change in rates. This indicates that the modified duration measure effectively overestimates the price sensitivity of lower coupon bonds. This factor is significant when hedging a mortgage-backed bond position, because using the modified duration figure to calculate the nominal value of the hedging instrument will not prove effective for anything other than very small changes in yield.

The formula to calculate approximate convexity (*or effective convexity*) is given below as (13.18); again if the values used in the formula allow for the cash flow to change, the convexity value may be taken to be the effective convexity. The effective convexity value of a mortgage pass-through security is invariably negative.

$$CV_{app} = \frac{P_+ + P_- - 2P_0}{P_0(\Delta rm)^2} \tag{13.18}$$

Selected bibliography and references

Babbel, D. and Zenios, S. 'Pitfalls in the analysis of option-adjusted spreads', *Financial Analysts Journa*l, July–Aug 1992, pp. 65–9.

Black, F., Derman, E. and Toy, W. 'A one-factor model of interest rates and its application to treasury bond options', *Financial Analysts Journal*, Jan–Feb 1990, pp. 33–9.

Boyce, W. and Koenigsberg, M. *Effective Duration of Callable Bonds: The Salomon Brothers Term Structure-Based Option Pricing Model*, Salomon Brothers Inc., April 1987.

Brown, S. *et al.*, *Analysis of Mortgage Servicing Portfolios,* Financial Strategies Group, Brudential-Bache Capital Funding, 1990.

Cox, J. C. Jr., Ingersoll, J. E. and Ross, S. A. 'A theory of the term structure of interest rates', *Econometrica*, March 1985, pp. 385–407.

Douglas Howard, C. 'Numerical pitfalls of lattice-based duration and convexity calculations', working paper, Department of Applied Mathematics and Physics, Polytechnic University, 1997.

Fabozzi, F. *Fixed Income Mathematics*, McGraw-Hill, 1997.

Fabozzi, F. (ed.) *Handbook of Structured Financial Products*, FJF Associates, 1998.

Fabozzi, F., Ramsey, C. and Ramirez, F. *Collateralized Mortgage Obligations: Structures and Analysis*, FJF Associates 1994, ch. 3.

Hayre, L., Mohebbi, C. and Zimmerman, T. 'Mortgage pass-through securities' in F. Fabozzi (ed.), *The Handbook of Fixed-Income Securities*, 5th edn, McGraw-Hill, 1997.

Herskovitz, M. D. 'Option-adjusted spread analysis for mortgage-backed securities', in F. Fabozzi (ed.), *The Handbook of Fixed-Income Options*, Probus, 1989, ch. 21.

Pinkus, S. and Chandoha, M. 'The relative price volatility of mortgage securities', in F. Fabozzi (ed), *Mortgage-Backed Securities: New Strategies, Applications and Research*, Probus, 1987.

Waldman, M. and Modzelewski, S. 'A framework for evaluating treasury-based adjustable rate mortgages', in Fabozzi, F. (ed.), *The Handbook of Mortgage-Backed Securities*, Probus, 1985.

14
Collateralised Debt Obligations

Collateralised debt obligations (CDOs) are a form of securitised debt. The market in such bonds emerged in the United States in the late 1980s, first as a form of repackaged high-yield bonds. The market experienced sharp growth in the second half of the 1990s, due to a combination of investor demand for higher yields allied to credit protection, and the varying requirements of originators, such as balance sheet management and lower-cost funding.

The term 'CDO' is a generic one, used to cover what are known as *collateralised bond obligations* (CBOs) and *collateralised loan obligations* (CLOs). Put simply, a CBO is an issue of rated securities backed or 'collateralised' by a pool of debt securities. A CLO on the other hand is an issue of paper that has been secured by a pool of bank loans. As the market has grown the distinction between the two types of structure has become blurred somewhat. Practitioners have taken to defining different issues in terms of the issuer's motivation, the type of asset backing and the type of market into which the paper is sold, for example the commercial paper market. Different structures are now more often categorised as being *balance sheet* transactions, usually a securitisation of assets in order to reduce regulatory capital requirements and provide the originator with an alternative source of funding, or *arbitrage* transactions, in which the originator sets up a managed investment vehicle in order to benefit from a funding gap that exists between assets and liabilities. Balance sheet transactions are issuer driven, whereas arbitrage transactions are typically investor driven.

In this chapter we provide an introduction to the most common CDO structures, as well as an overview of the analysis of these instruments. We also provide an update for the market following the crash in 2007–2008.

An overview of CDOs

Introduction

Collateralised debt obligation or CDO is the generic term for two distinct products, so-called balance sheet transactions and arbitrage transactions. The common thread between these structures is that they are both backed by some form of commercial or corporate debt or loan receivable. The primary differences between the two types are the type of collateral backing the newly created securities in the CDO

structure, and the motivations behind the transaction. The growth of the market has been in response to two key requirements: the desire of investors for higher-yield investments in higher-risk markets, managed by portfolio managers skilled at extracting value out of poorly performing or distressed debt, and the need for banks to extract greater value out of assets on their balance sheet, almost invariably because they are generating a below-market rate of return. By securitising bond or loan portfolios, banks can lower their capital charge by removing them from their balance sheet and funding them at a lower rate. The market has its origins in investor-driven arbitrage transactions, with bank balance sheet transactions a natural progression after banks applied securitisation techniques to their own asset base.

Balance sheet CDOs are structured securities that are usually backed with bank-originated, investment grade commercial and corporate loans. Since this form of collateral is almost invariably loans, and very rarely bonds, these transactions are usually referred to as CLOs. Why would a bank wish to securitise part of its loan portfolio? In short, it does so in order to improve its capital adequacy position. Securitising a bank's loans reduces the size of its balance sheet, thereby improving its capital ratio and lowering its capital charge. The first domestic balance sheet CLO in the US market was the NationsBank Commercial Loan Master Trust, series 1997-1 and 1997-2, issued in September 1997, which employed what is known as a Master Trust structure to target investors who had previously purchased asset-backed securities (ABS).[1] As balance sheet CLOs are originated mainly by commercial banks, the underlying collateral is usually part of their own commercial loan portfolios, and can be fixed term, revolving, secured and unsecured, syndicated and other loans. Although we note that most CLOs have been issued by banks that are domiciled in the main developed economies, geographically the underlying collateral often has little connection with the home country of the originating bank. Most bank CLOs are floating-rate loans with average lives of five years or less. They are targeted mainly at bank sector Libor-based investors, and are structured with an amortising payoff schedule.

Arbitrage CDOs are backed with high-yield corporate bonds or loans. As the collateral can take either form, arbitrage CDOs can be either CLOs or CBOs. Market practitioners often refer to all arbitrage deals as CDOs for simplicity, irrespective of the collateral backing them. The key motivation behind arbitrage CDOs is, unsurprisingly, the opportunity for arbitrage, or the difference between investment-grade funding rates and high-yield investment rates. In an arbitrage CDO, the income generated by the high-yield assets should exceed the cost of funding, as long as no credit event or market event takes place.

[1] The master trust structure is a generic set up that allows originators to issue subsequent asset-backed deals under the same legal arrangement, thus enabling such issues to be made quicker than they otherwise might be. Investors also welcome such a structure, as they indicate a commitment to liquidity by implying further issues into the market.

Security structure

Arbitrage and balance sheet CDOs generally have similar structures. In essence a special purpose vehicle (SPV) purchases loans or bonds directly from the originator or from the secondary market. SPVs have been set up in a number of ways, which include *special purpose corporations*, limited partnerships and limited liability corporations. An SPV will be bankruptcy-remote, that is, unconnected to any other entities that support it or are involved with it. The parties involved in a transaction, apart from the investors, are usually a portfolio manager, a bond trustee appointed to look after the interests of the investors, a credit enhancer and a back-up servicer. Some structures involve a swap arrangement where this is required to alter cash flows or set up a hedge, so in such cases a swap counterparty is also involved. A basic structure is shown in Figure 14.2, which is applicable to both CBOs and CLOs.

Invariably the SPV will issue a number of classes of debt, with credit enhancements included in the structure such that different tranches of security can be issued, each with differing levels of credit quality. Other forms of protection may also be included, including a cash reserve.

A further development has been the issue of *synthetic* CDOs. Synthetic CLO structures use credit derivatives that allow the originating bank to transfer the risk of the loan portfolio to the market. In a synthetic structure there is no actual transfer of the underlying reference assets; instead the economic effect of a traditional CDO is synthesised by passing to the end investor(s) an identical economic risk to that associated with the underlying assets that would have been transferred for a conventional CDO. This effect is achieved by the provision by a counterparty of a *credit default swap*, or the issue of credit-linked notes by the originating bank, or a combination of these approaches. In a credit-linked CLO the loan portfolio remains on the sponsoring bank's balance sheet, and investors in the securities are exposed to the credit risk of the bank itself, in addition to the market risk of the collateralised portfolio. Therefore the credit rating of the CLO can be no higher than that of the originating bank.

Comparisons with other asset-backed securities

The CDO asset class has similarities in its fundamental structure with other securities in the ABS market. Like other ABS, a CDO is a debt obligation issued by an SPV, secured by a form of receivable. In this case, though, the collateral concerned is high-yield loans or bonds, rather than, for example, mortgage or credit card receivables. Again similar to other ABS, CDO securities typically consist of different credit tranches within a single structure, and the credit ratings range from AAA to B or unrated. The rating of each CDO class is determined by the amount of credit enhancement in the structure, the ongoing performance of the collateral, and the priority of interest in the cash flows generated by the pool of assets.

The credit enhancement in a structure is among items scrutinised by investors, who will determine the cash flow *waterfalls* for the interest and principal, the prepayment conditions, and the methods of allocation for default and recovery. Note that the term 'waterfall' is used in the context of asset-backed securitisations that

are structured with more than one tranche, to refer to the allocation of principal and interest to each tranche in a series. If there is excess cash and this can be shared with other series, the cash flows are allocated back through the waterfall, running over the successive tranches in the order of priority determined at issue.

A significant difference between CDOs and other ABS is the relationship with the servicer. In a traditional ABS the servicing function is usually performed by the same entity that sources and underwrites the original loans. These roles are different in a CDO transaction; for instance there is no servicer that can collect on non-performing loans. Instead the portfolio manager for the issuer must actively manage the portfolio. This might include sourcing higher-quality credits, selling positions before they deteriorate, and purchasing investments that are expected to appreciate. In essence portfolio managers assume the responsibility of a servicer. Therefore investors in CDOs must focus their analysis on the portfolio manager as well as on the credit quality of the collateral pool. CDO structures also differ from other ABS in that they frequently hold non-investment grade collateral in the pool, which is not a common occurrence in traditional ABS structures. Finally CDO transactions are (or rather, have been to date) private and not public securities.

CDO asset types

The arbitrage CDO market can be broken down into two main asset types, *cash flow* and *market value* CDOs. Balance sheet CDOs are all cash flow CDOs.

Cash flow CDOs share more similarities with traditional ABS than do market value transactions. Collateral is usually a self-amortising pool of high-yield bonds and loans, expected to make principal and interest payments on a regular basis. Most cash flow CDO structures allow for a reinvestment period, and while this is common in other types of ABS, the period length tends to be longer in cash flow CDOs, typically with a minimum of four years. The cash flow structure relies upon the collateral's ability to generate sufficient cash to pay principal and interest on the rated classes of securities. This is similar to an automobile ABS, in which the auto-backed securities rely upon the cash flows from the fixed pool of automobile loans to make principal and interest payments on the liabilities. Trading of the CDO collateral is usually limited, for instance in the event of a change in credit situation, and so the value of the portfolio is based on the par amount of the collateral securities.

Market value CDOs, which were first introduced in 1995, resemble hedge funds more than traditional ABS. The main difference between a cash flow CDO and a market value CDO is that the portfolio manager has the ability to freely trade the collateral. This means investors focus on expected appreciation in the portfolio, and the portfolio itself may be quite different in, say, three months' time. This leads to the analogy with the hedge fund. Investors in market value CDOs are as concerned with the management and credit skills of the portfolio manager as they are with the credit quality of the collateral pool. Market value CDOs rely upon the portfolio manager's ability to generate total returns and liquidate the collateral in timely fashion, if necessary, in order to meet the cash flow obligations (principal and interest) of the CDO structure.

Different portfolio objectives result in distinct investment characteristics. Cash flow CDO assets consist mainly of rated, high-yield debt or loans that are *current* in their principal and interest payments, that is, they are not in default. In a market value CDO the asset composition is more diversified. The collateral pool might consist of, say, a 75:25 percentage split between assets to support liability payments and investments to produce increased equity returns. In this case, the first 75% of assets of a market value CDO will resemble those of a conventional cash flow CDO, with say 25% invested in high-yield bonds and 50% in high-yield loans. These assets should be sufficient to support payments on 100% of the liabilities. The remaining 25% of the portfolio might be invested in 'special situations' such as distressed debt, foreign bank loans, hybrid capital instruments and other investments. The higher-yielding investments are required to produce the higher yields that are marketed to equity investors in market value CDOs.

We have described in general terms the asset side of a CDO. The liability side of a CDO structure is similar to other ABS structures, and encompasses several investment grade and non-investment grade classes with an accompanying equity tranche that serves as the first loss position. In for example a mortgage-backed transaction, the equity class is not usually offered but instead is held by the issuer. Typically in the US market, rated CDO liabilities have a 10–12-year legal final maturity. The four main rating agencies[2] all actively rate cash flow CDOs, although commonly transactions carry ratings from only one or two of the agencies.

Liabilities for market value CDOs differ in some ways from cash flow CDOs. In most cases senior bank facilities provide more than half of the capital structure, with a six to seven-year final maturity. When a market value transaction is issued, cash generated by the issuance is usually not fully invested at the start. There is a *ramp-up* period to allow the portfolio manager time to make investment decisions and effect collateral purchases. Ramp-up periods result in a risk that cash flows on the portfolio's assets will not be sufficient to cover liability obligations at the start. Rating agencies consider this ramp-up risk when evaluating the transaction's credit enhancement. Ramp-up periods are in fact common to both cash flow and market value CDOs, but the period is longer with the latter transactions, resulting in more significant risk.

Investor analysis

CDOs offer investors a variety of risk/return profiles, as well as market volatilities, and their appeal has widened as broader macroeconomic developments in the global capital markets have resulted in lower yields on more traditional investments.

Investors analysing CDO instruments will focus on particular aspects of the market. For instance those with a low appetite for risk will concentrate on the higher-rated classes of cash flow transactions. Investors that are satisfied with greater volatility of

[2] That is, Standard & Poor's, Moody's and Fitch.

earnings but still wish to hold AA or AAA-rated instruments may consider market value deals. The 'arbitrage' that exists in the transaction may be a result of:

- industry diversification
- differences between investment grade and high-yield spreads
- the difference between implied default rates in the high-yield market and expected default rates
- the liquidity premium embedded in high-yield investments
- the Libor rate versus the Treasury spread.

The CDO asset class cannot be compared in a straightforward fashion with other ABS classes, which makes relative value analysis difficult. Although a CDO is a structured finance product, it does not have sufficient common characteristics with other such products. The structure and cash flow of a CDO are perhaps most similar to a commercial mortgage-backed security (MBS); the collateral backing the two types shares comparable characteristics. Commercial mortgage pools and high-yield bonds and loans both have fewer obligors and larger balances than other ABS collateral, and each credit is rated. On the other hand CDOs often pay floating-rate interest and are private securities,[3] whereas commercial MBS (in the US market) pay fixed rate and are often public securities.

We have presented only an overview here; interested readers may wish to consult the references at the end of this chapter, principally Fabozzi (1998, ch. 18) and the study by ING Barings.

Cash flow CDO structures

As we noted earlier, cash flow CDOs are categorised as either balance sheet or arbitrage deals. Arbitrage CDOs are further categorised as cash flow or market-value deals. The total issuance of all cash flow deals in 2000 and 2001 is shown in Figure 14.1.

Cash flow CDOs are similar to other asset-backed securitisations involving an SPV. Bonds or loans are pooled together and the cash flows from these assets are used to back the liabilities of the notes issued by the SPV into the market. As the underlying assets are sold to the SPV, they are removed from the originator's balance sheet; hence the credit risk associated with these assets is transferred to the holders of the issued notes. The originator also obtains funding by issuing the notes.

Banks and other financial institutions are the primary originators of balance sheet CDOs. These are deals securitising banking assets such as commercial loans

[3] In the US market, they are also filed under Rule 144A, as opposed to public securities which must be registered with the Securities and Exchange Commission. Rule 144A securities may only be sold to investors classified as professional investors under specified criteria. Rule 144A provides an exemption from the registration requirements of the Securities Act (1933) for resale of privately placed securities to qualified institutional buyers. Such buyers are deemed to be established and experienced institutions, and so the SEC does not regulate or approve disclosure requirements.

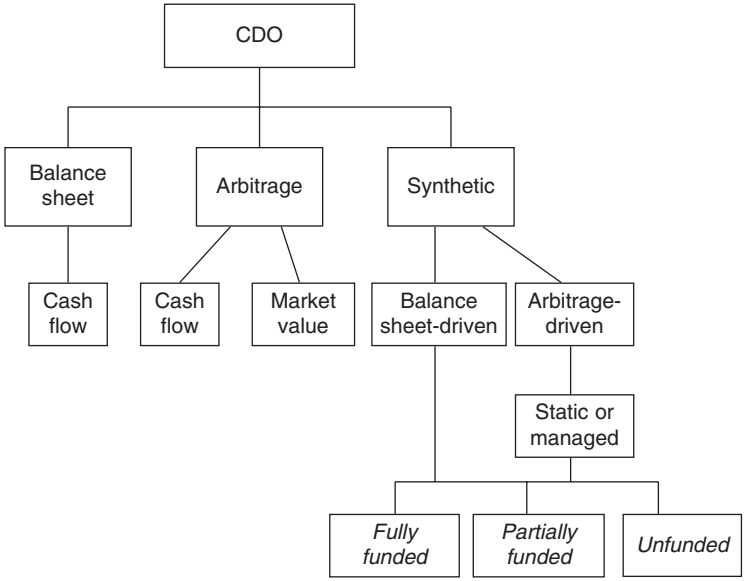


Figure 14.1 The CDO family

of investment grade or sub-investment grade rating. The main motivations for entering into this arrangement are:

- to obtain regulatory relief
- to increase return on capital via the removal of lower yielding assets from the balance sheet
- to secure alternative and/or cheaper sources of funding.

Investors are often attracted to balance sheet CDOs because they are perceived as offering a higher return than, for example, credit card ABS at a similar level of risk exposure. They also represent a diversification away from traditional structured finance investments. The asset pool in a balance sheet CDO is static, that is, it is not traded or actively managed by a portfolio manager. For this reason the structure is similar to more traditional ABS or repackaging vehicles. The typical note tranching is:

- senior note, AAA-rated, and 90–95% of the issue
- subordinated note, A-rated, 3–5%
- mezzanine note, BBB-rated, 1–3%
- equity note, non-rated, 1–2%.

The cash flows of the underlying assets are used to fund the liabilities of the overlying notes. As the notes carry different ratings, there is a *priority of payment* that must be followed, the cash flow waterfall mentioned earlier.

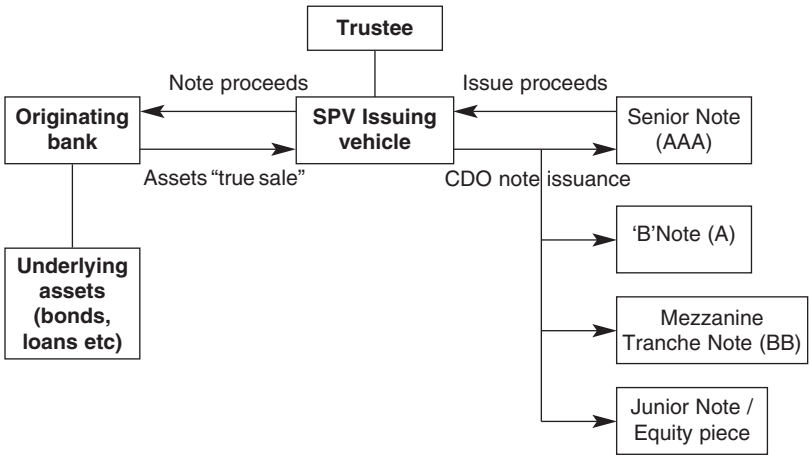


Figure 14.2 Generic cash flow CDO

The waterfall process for interest payments is shown in Figure 14.3. Before paying the next priority of the waterfall, the vehicle must pass a number of compliance tests on the level of its underlying cash flows. These include interest coverage and principal (par) coverage tests.

During the life of the CDO transaction, a portfolio administrator will produce a periodic report detailing the quality of the collateral pool. This report is known as an *investor or trustee report* and also shows the results of the compliance tests that are required to affirm that the notes of the CDO have maintained their credit rating.

As noted earlier, arbitrage CDOs are classified into either cash flow CDOs or market value CDOs, with the designation depending on the way the underlying asset pool is structured to provide value (cash proceeds) in the vehicle. The distinction is that:

- A cash flow CDO will have a collateral pool that is usually static, and which generates sufficient interest to cover fees, expenses and overlying liabilities, and sufficient principal to repay notes on maturity.
- In a market value CDO the collateral pool is usually actively traded, and marked-to-market on a daily basis. The daily mark for all the assets indicates the total value of the collateral pool, and this value must always be sufficient to cover the principal and interest liabilities.

A cash flow arbitrage CDO has certain similarities with a balance sheet CDO, and if it is a static pool CDO it is also conceptually similar to an ABS deal.[4] The

[4] Except that in a typical ABS deal such as a consumer or trade receivables deal, or a residential MBS deal, there are a large number of individual underlying assets, whereas with a CBO or CLO there may be as few as 20 underlying loans or bonds.

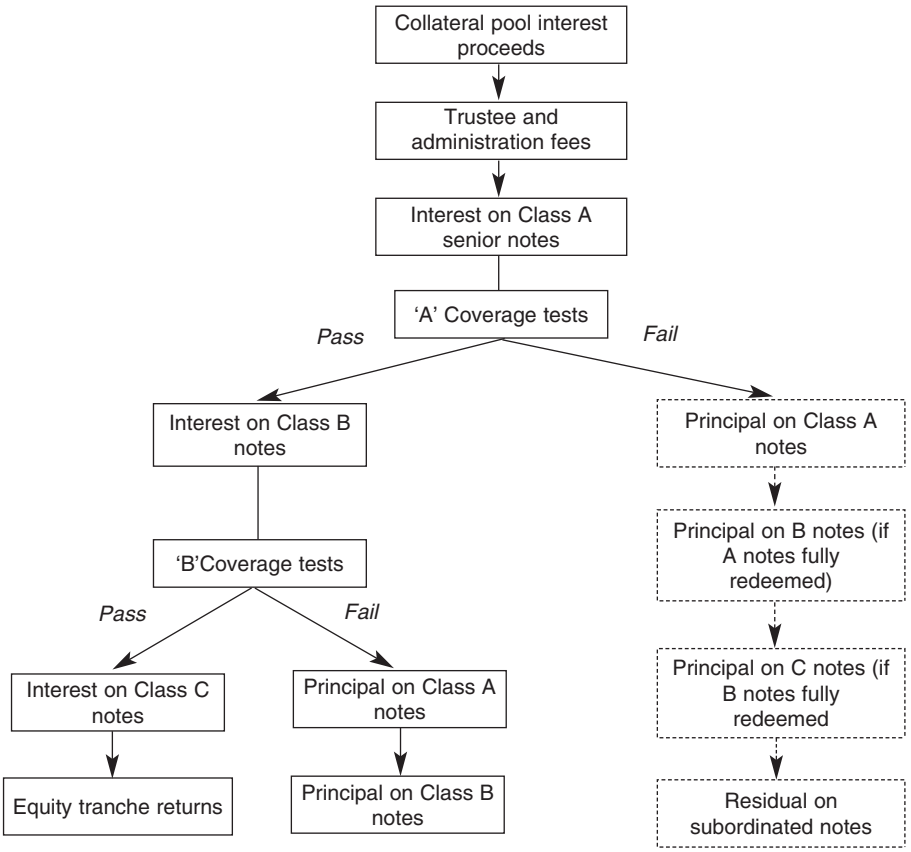


Figure 14.3 Interest cash flow waterfall for cash flow CDO

priority of payments is similar, starting from expenses, trustee and servicing fees, senior noteholders, and so on down to the most junior noteholder. Underlying assets on cash flow arbitrage deals are commonly lower-rated bonds, commercial bank loans, high-yield debt and emerging market sovereign bonds. The basic structure is designed to split the aggregate credit risk of the collateral pool into various tranches, which are the overlying notes, each of which has a different credit exposure from the other. As a result each note figures a different risk/reward profile, and so will attract different classes of investor.

The issued notes have different risk profiles because they are subordinated, that is, the notes are structured in descending order of seniority. In addition the structure makes use of credit enhancements to varying degrees, which include:

- *Over-collateralisation*: the overlying notes are lower in value than the underlying pool. For example, $250 million nominal of assets are used as backing for $170 million nominal of issued bonds.

- *Cash reserve accounts*: a reserve is maintained in a cash account and used to cover initial losses. The funds may be sourced from part of the proceeds.
- Excess spread: cash inflows from assets that exceed the interest service requirements of liabilities.
- *Insurance wraps*: insurance cover against losses suffered by the asset pool, for which an insurance premium is paid for as long as the cover is needed.

The quality of the collateral pool is monitored regularly and reported on by the portfolio administrator, who produces the investor report. This report details the results of various compliance tests, which are undertaken at individual asset level as well as aggregate level. Compliance tests include:

- *Weighted average spread and weighted average rating*: the average interest spread and average credit rating of the assets, which must remain at a specified minimum.
- *Concentration*: there will be a set maximum share of the assets that may be sourced from particular emerging markets, industrial sectors, and so on.
- *Diversity score*: this is a statistical value that is calculated via a formula set by the rating agency analysing the transaction. It measures the level of diversity of the assets, in other words how different they are – and hence how uncorrelated in their probability of default – from each other.

These tests are calculated on a regular basis and also each time the composition of the assets changes, for example because certain assets have been sold, new assets have been purchased, or bonds have paid off ahead of their legal maturity date. If the test results fall below the required minimum, trading activity is restricted to only those trades that will improve the test results. Certain other compliance tests are viewed as more important, since if any of them are 'failed', the cash flows will be diverted from the normal waterfall and will be used to begin paying off the senior notes until the test results improve. These include:

- *Over-collateralisation*: the over-collateralisation level for the issued notes must remain above a specified minimum. For instance it must be at 120% of the nominal value of the senior note.
- *Interest coverage*: the level of interest receivables on assets must be sufficient to cover interest liabilities but also to bear default and other losses.

Compliance tests are specified as part of the process leading up to the issue of notes, in discussion between the originator and the rating agency.[5] The ratings analysis is comprehensive, and focuses on the quality of the collateral, individual asset default probabilities, the structure of the deal and the track record and reputation of the originator.

[5] Deals may be rated by more than one rating agency.

Synthetic CDOs

Compared with conventional cash flow deals, which feature an actual transfer of ownership or true sale of the underlying assets to a separately incorporated legal entity, a synthetic securitisation structure is engineered so that the credit risk of the assets is transferred by the sponsor or originator of the transaction, from itself, to the investors by means of credit derivative instruments. The originator is therefore the credit protection buyer and investors are the credit protection sellers. This credit risk transfer may be undertaken either directly or via an SPV. Using this approach, underlying or *reference* assets are not necessarily moved off the originator's balance sheet, so the approach is adopted whenever the primary objective is to achieve risk transfer rather than balance sheet funding. The synthetic structure enables removal of credit exposure without asset transfer, so may be preferred for risk management and regulatory capital relief purposes. For banking institutions it also enables loan risk to be transferred without selling the loans themselves, thereby allowing customer relationships to remain unaffected. The first synthetic deals were observed in the US market, while the first deals in Europe were observed in 1998.

The first European synthetic deals were balance sheet CLOs, with underlying reference assets being commercial loans on the originator's balance sheet. Originators were typically banking institutions. Arbitrage synthetic CDOs have also been introduced, typically by fund management institutions, and involve sourcing credit derivative contracts in the market and then selling these on to investors in the form of rated notes, at the arbitrage profit. Within the synthetic market, arbitrage deals were the most frequently issued during 2001, reflecting certain advantages they possess over cash CDOs. A key advantage has been that credit default swaps for single reference entities frequently trade at a lower spread than cash bonds of the same name and maturity, with consequently lower costs for the originator.

Motivations

The differences between synthetic and cash CDOs are perhaps best reflected in the different cost–benefit economics of issuing each type. The motivations behind the issue of each type usually also differ. A synthetic CDO can be seen as being constructed out of the following:

- a short position in a credit default swap (bought protection), by which the sponsor transfers its portfolio credit risk to the issuer (SPV)
- a long position in a series of tranched liabilities of overlying notes or CDS, issued by the SPV.

The originators of the first synthetic deals were banks that wished to manage the credit risk exposure of their loan books, without having to resort to the administrative burden of true sale cash securitisation. They are a natural progression in the development of credit derivative structures, with single name credit default swaps being replaced by portfolio default swaps. Synthetic CDOs can be 'de-linked' from

the sponsoring institution, so that investors do not have any credit exposure to the sponsor itself. The first deals were introduced (in 1998) at a time when widening credit spreads and the worsening of credit quality among originating firms meant that investors were sellers of cash CDOs which had retained a credit linkage to the sponsor. A synthetic arrangement also means that the credit risk of assets that are otherwise not suited to conventional securitisation may be transferred, while assets are retained on the balance sheet. Such assets include bank guarantees, letters of credit or cash loans that have some legal or other restriction on being securitised. For this reason synthetic deals are more appropriate for assets that are described under multiple legal jurisdictions.

The economic advantage of issuing a synthetic versus a cash CDO can be significant. Put simply, the net benefit to the originator is the gain in regulatory capital cost, minus the cost of paying for credit protection on the credit default swap side. In a partially funded structure, a sponsoring bank will obtain full capital relief when note proceeds are invested in 0% risk-weighted collateral such as Treasuries or gilts. The super senior swap portion will carry a 20% risk weighting.[6]

Synthetic deals can be unfunded, partially funded or fully funded. An unfunded CDO would be comprised wholly of credit default swaps, while fully funded structures would be arranged so that the entire credit risk of the reference portfolio was transferred through the issue of credit-linked notes. We discuss these shortly.

Mechanics

A synthetic CDO is so called because the transfer of credit risk is achieved 'synthetically' via a credit derivative, rather than by a 'true sale' to an SPV. Thus in a synthetic CDO the credit risk of the underlying loans or bonds is transferred to the SPV using credit default swaps and/or total return swaps (TRS). However the assets themselves are not legally transferred to the SPV, and they remain on the originator's balance sheet. Using a synthetic CDO, the originator can obtain regulatory capital relief[7] and manage the credit risk on its balance sheet, but will not be receiving any funding. In other words a synthetic CDO structure enables originators to separate credit risk exposure and asset funding requirements. The credit risk of the asset portfolio, now known as the *reference portfolio*, is transferred, directly or to an SPV, through credit derivatives. The most common credit contracts used are credit default swaps. A portion of the credit risk may be sold on as credit-linked notes. Typically a large majority of the credit risk is transferred via a 'super-senior' credit default swap,[8] which is dealt with a swap counterparty but usually sold to monoline insurance companies at a significantly lower spread over Libor than the senior AAA-rated tranche of cash flow CDOs. This is a key attraction of synthetic deals for originators.

[6] This is as long as the counterparty is an OECD bank, which is invariably the case.
[7] This is because reference assets that are protected by credit derivative contracts, and which remain on the balance sheet will, under Basel rules, attract a lower regulatory capital charge.
[8] So called because the swap is ahead of the most senior of any funded (note) portion, which latter being 'senior' means the swap must be 'super-senior'

Most deals are structured with mezzanine notes sold to a wider set of investors, the proceeds of which are invested in risk-free collateral such as Treasury bonds or Pfandbriefe securities. The most junior note, known as the 'first loss' piece, may be retained by the originator. On occurrence of a credit event among the reference assets, the originating bank receives funds remaining from the collateral after they have been used to pay the principal on the issued notes, less the value of the junior note.

A generic synthetic CDO structure is shown in Figure 14.4. In this generic structure, the credit risk of the reference assets is transferred to the issuer SPV and ultimately the investors, by means of the credit default swap and an issue of credit-linked notes. In the default swap arrangement, the risk transfer is undertaken in return for the swap premium, which is then paid to investors by the issuer. The note issue is invested in risk-free collateral rather than passed on to the originator, in order to de-link the credit ratings of the notes from the rating of the originator.

If the collateral pool was not established, a downgrade of the sponsor could result in a downgrade of the issued notes. Investors in the notes expose themselves to the credit risk of the reference assets, and if there are no credit events they will earn returns at least the equal of the collateral assets and the default swap premium. If the notes are credit-linked, they will also earn excess returns based on the performance of the reference portfolio. If there are credit events, the issuer will

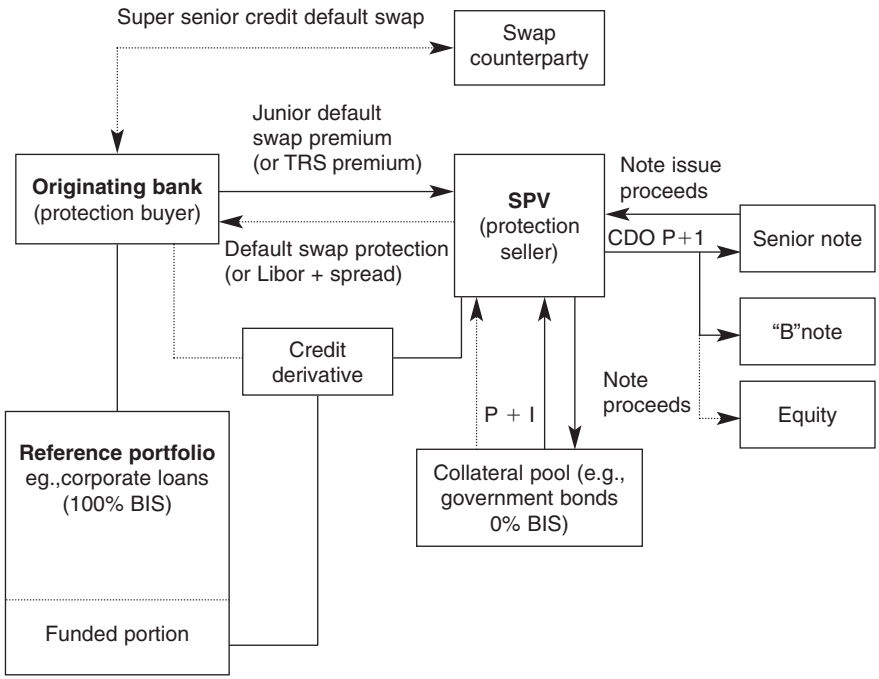


Figure 14.4 Synthetic CDO structure

deliver the assets to the swap counterparty and will pay the nominal value of the assets to the originator out of the collateral pool. *Credit default swaps* are unfunded credit derivatives, while CLNs are funded credit derivatives where the protection seller (the investors) funds the value of the reference assets up-front, and will receive a reduced return on occurrence of a credit event.

Synthetic CDOs were popular in the European market because of their less onerous legal documentation requirements. Across Europe the legal infrastructure is not uniform, and in certain countries it has not been sufficiently developed to enable true sale securitisation to be undertaken. In addition when the underlying asset pool is composed of bonds from different countries, a cash funded CDO may present too many administrative difficulties. A synthetic CDO removes such issues by using credit derivatives, and in theory can be brought to market more quickly than a cash flow arrangement (although in practice this is not always the case). It is also able to benefit from a shorter ramp-up period prior to closing.

Traditional cash CDOs suffer when only a portion of the portfolio is in place at time of closing, especially holders of the equity piece who may receive no cash inflow during the ramp-up. This cost may be reduced if a *warehouse* facility is set up prior to closing the transaction. A warehouse, which is provided by the portfolio administrator in return for a fee, enables the CDO issuer to begin purchasing bonds before closing. Using a warehouse arrangement, the issuer purchases target securities and hedges their interest-rate risk using futures, interest-rate swaps or by shorting government bonds. The credit risk on the bonds is not hedged, of course, and equity holders bear losses if spreads widen due to credit deterioration. (They will of course gain if spreads tighten.) A shorter ramp-up period is a plus therefore, and although in a synthetic deal the manager will take some time while sourcing credit default swaps, this time is typically shorter than in a cash flow CDO.

Within the European market, static synthetic balance sheet CDOs are the most common structure. There are two reasons that banks originate them:

- *Capital relief*. Banks can obtain regulatory capital relief by transferring lower-yield corporate credit risk such as corporate bank loans off their balance sheet. Under Basel I rules all corporate debt carried an identical 100% risk-weighting; therefore with banks having to assign 8% of capital for such loans, higher-rated (and hence lower-yielding) corporate assets will require the same amount of capital but will generate a lower return on that capital. A bank may wish to transfer such higher-rated, lower-yield assets from its balance sheet, and this can be achieved via a CDO transaction. The capital requirements for a synthetic CDO are lower than for corporate assets; for example the funded segment of the deal will be supported by high-quality collateral such as government bonds, and arranged via a repo with an OECD bank which would carry a 20% risk weighting. This is the same for the super senior element.
- *Transfer of credit risk*. The cost of servicing a fully funded CDO, and the premium payable on the associated credit default swap, can be prohibitive. With a partially funded structure, the issue amount is typically a relatively small share of the asset portfolio. This lowers substantially the default swap premium. Also,

as the CDO investors suffer the first loss element of the portfolio, the super senior default swap can be entered into at a considerably lower cost than that on a fully funded CDO.

Synthetic deals may be either static or managed. Static deals have two advantages: there are no ongoing management fees to be borne by the vehicle, and the investor can review and grant approval to credits that are to make up the reference portfolio. The disadvantage is that if there is a deterioration in credit quality of one or more names, there is no ability to remove or offset this name from the pool, and the vehicle continues to suffer from it.

Risk and return on CDOs

Risk–return analysis

The return analysis for CDOs performed by potential investors is necessarily different from that undertaken for other securitised asset classes. For CDOs the three key factors to consider are:

- default probabilities and cumulative default rates
- default correlations
- recovery rates.

Analysts make assumptions about each of these with regard to individual reference assets, usually with recourse to historical data. We consider each factor in turn.

Default probability rates

The level of default probability rates will vary with each deal. Analysts such as the rating agencies will use a number of methods to estimate default probabilities, such as individual reference credit ratings and historical probability rates. Since there may be as many as 150 or more reference names in the CDO, a common approach is to use the average rating of the reference portfolio. Rating agencies such as Moody's provide data on the default rates for different ratings as an 'average' class, which can be used in the analysis.

Correlation

The correlation between assets in the reference portfolio of a CDO is an important factor in CDO returns analysis. A problem arises with what precise correlation value to use: these can be correlation between default probabilities, correlation between timing of default and correlation between spreads. The *diversity score* value of the CDO plays a part in this: it represents the number of uncorrelated bonds with identical par value and the same default probability.

Recovery rates

Recovery rates for individual obligors differ by issuer and industry classification. Rating agencies such as Moody's publish data on the average prices of all defaulted

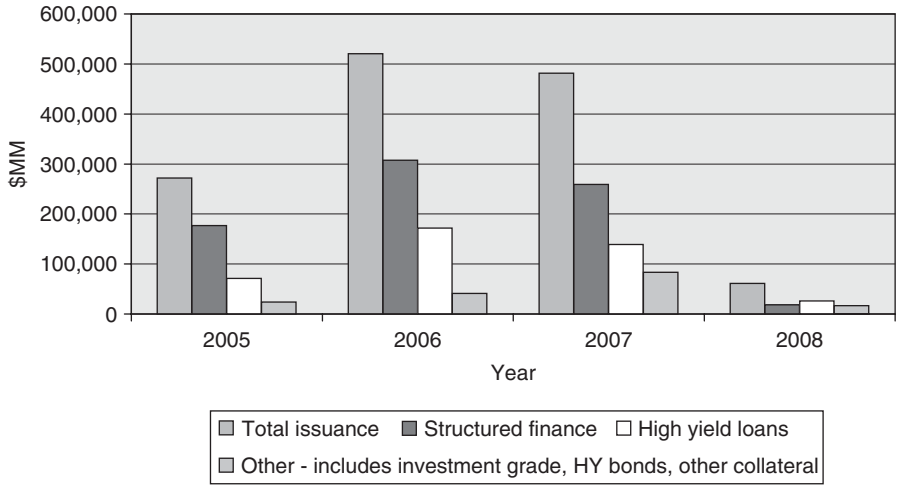


Figure 14.5 Total issuance by CDO underlying
Source: SIFMA

bonds, and generally analysts will construct a database of recovery rates by industry and credit rating for use in modelling the expected recovery rates of assets in the collateral pool. Note that for synthetic CDOs with credit default swaps as assets in the portfolio, this factor is not relevant.

Analysts undertake simulation modelling to generate scenarios of default and expected return. For instance they may model the number of defaults up to maturity, the recovery rates of these defaults and the timing of defaults. All these variables are viewed as random variables, so they are modelled using a stochastic process.

CDO market overview since 2005

The period 2005 to the first half of 2007 was a period of very high growth for the issuance of CDOs. During this period the growth in the CDO product was investor-driven and high volumes of arbitrage CDO transactions were issued. However, a big impact of the credit crunch from 2007 has been a sharp decline in new issuance of CDO product.

Figures 14.5 and 14.6 illustrate CDO issuance volumes during 2005–2008, by collateral type. The impact of the credit crunch on issuance volumes is plain to see.

The growth in the CDO market had been driven by large volumes of ABS CDO (collateralised debt obligations referencing ABS) and CLO (collateralised loan obligations):

The ABS CDO product is typically divided into 2 groups:

(i) high-grade ABS CDO: collateral is ABS originally rated in the range AAA to A or
(ii) mezzanine ABS CDO: collateral is ABS originally rated in the range BBB.

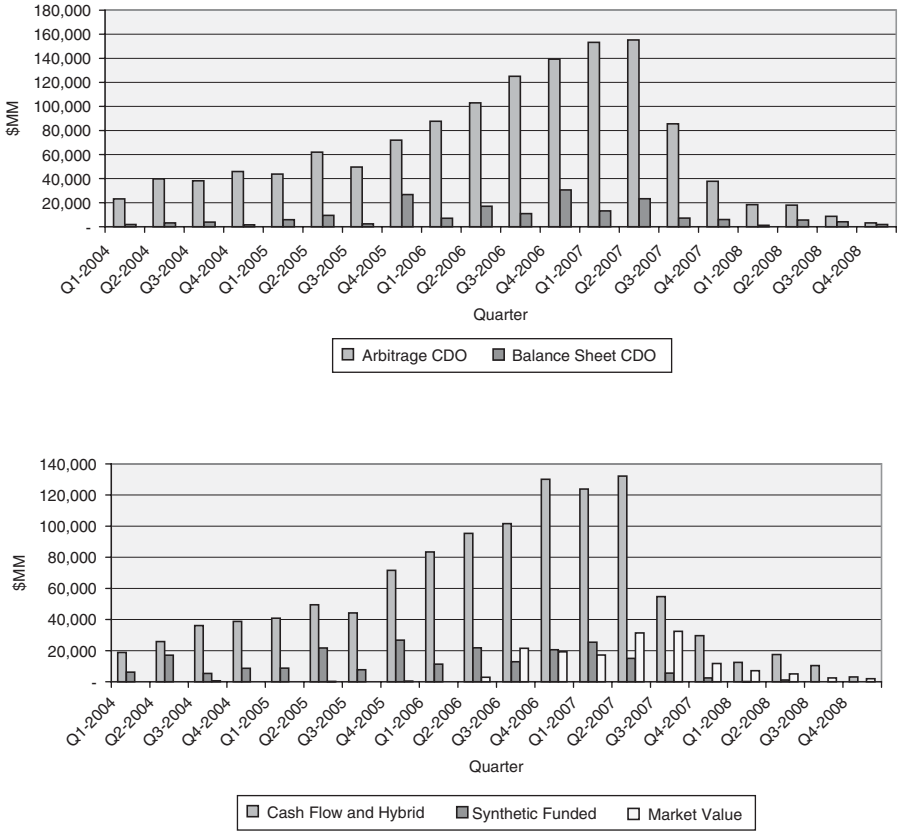


Figure 14.6 Issuance data volumes, arbitrage CDO and balance sheet CDO
Source: SIFMA

The investor-related weakness of the ABS CDO product was its significant concentration in newly issued US sub-prime RMBS which subsequently performed poorly from 2006 onwards, mainly due to poor origination standards and rising USD interest rates.

The credit crunch resulted in some large banks announcing significant write downs on super-senior ABS CDO positions that were held on their balance sheets and had originally been rated AAA.

By the end of 2007 the growth in volumes of CLO issuance had become so significant, that CLO vehicles were the largest non-bank lenders or purchasers of leveraged loans in the primary market. These leveraged loans were used in leveraged buyouts (LBOs) and formed an important source of financing to the private equity sector.

Since the start of the credit crunch there has been a sharp decline in the issuance volumes for both ABS CDO and CLO in 2008, as evident from Figure 14.6.

Risk and capital management

At the height of the CDO market, during the years 2005 to mid-2007, the financial system suffered from weak risk and capital management in a number of areas that were related to the CDO product.

Subprime origination and underwriting standards

The underwriting standards fell significantly in the years 2005 to 2007 as subprime mortgages were financed and included in US RMBS collateral pools. The ABS CDO was a major investor in US subprime RMBS. Both US RMBS and ABS CDO performance suffered as a result of the poor underwriting in subprime mortgages.

Failure of hedge counterparties ('wrong way exposure')

Banks tried to hedge their exposure to super-senior ABS CDO by buying credit protection from monoline insurers. The typical transaction involved a negative basis trade in which the bank would hold the super-senior ABS CDO position, typically rated AAA and also buy credit protection from a monoline insurer. However, these hedges suffered from 'wrong way' exposure of the monoline counterparty. The 'wrong way' exposure exists because the counterparty's creditworthiness is lowest when the exposure is largest.

Reliance on credit rating agencies (CRAs)

The CRAs play a key role in the operation of credit markets and their rating is an indicator of credit quality. However, the rating is an opinion based on the CRA's methodology. The CRAs usually make it clear that their credit rating is a measure of credit quality, not an investment recommendation – and does not capture the risk of a decline in market value or liquidity of the product. It remained important for investors and originators to do significant amounts of their own independent credit analysis and due diligence, and assess the market risk and liquidity risk associated with an instrument, rather than rely purely on CRAs' rating. Unfortunately many investors did not perform such due diligence.

Poor risk assessment of the credit risks of warehoused assets and balance sheet positions

Banks were exposed to the performance of credit assets that were held on balance sheet awaiting securitisation. As the CDO issuance market started to fall in volume in mid-2007, banks were forced to write down some asset positions, as it became increasingly difficult to sell these assets to CDO vehicles and into the wider market. Many banks were positioned the same way and this caused a fall in the appetite to make bids for assets.

Some banks also built up large positions in super-senior ABS CDO and failed to capture the risk of these securities. These banks' objective was to hold these securities, and not distribute these risks. Some positions that were 'hedged' suffered from wrong way counterparty exposure risks.

Valuation and risk measurement

Positions that had been held on the trading books of some banks became difficult to value and risk manage, due to the withdrawal of liquidity. Some risk measurement systems did not make sufficient allowance for the risk of significant credit spread widening and credit migration, since they referred to risk data sets for the previous two to three years before 2007, which showed a fairly stable credit environment.

As a result there was an underestimation of the need for capital against credit positions. The key areas for improvement include: stress testing, scenario analysis and use of a risk framework that can assess the interaction between risks (i.e. liquidity, market and credit risk). Market participants should establish a validation process for models and pricing methodologies that can be used to risk manage exposures under stressed conditions. The valuation and risk management of 'two-layer' securitisations, that is ABS CDO, proved to be a significant weakness.

Correlation risk

The CDO product makes assumptions about the extent of diversification in the underlying asset or reference portfolio. In a stressed economy the ABS CDO product showed a significant exposure to systematic shocks (due to its exposure to the housing market via US RMBS). At issuance there had been high reliance on rating agency data, but the short data history on underlying subprime RMBS proved to be inadequate to manage the significant correlation risk.

Liquidity risk

The impact of falling demand for assets such as ABS, CDOs and leveraged loans during the second half of 2007 and 2008 due to the reduction in investor appetite caused a further problem: high liquidity risk. Structured investment vehicles and conduits with exposure to ABS investments, relied on rolling forward short-term commercial paper financing; however, these vehicles suffered from the withdrawal of short-term commercial paper financing by investors, who had a 'flight to quality' of short-dated government bonds. This created further liquidity strains for banks, as they tried to support vehicles they had sponsored to reduce their reputational risk and also to honour liquidity facilities.

Structured finance became an integral part of the banking and financial system in the years before the credit crunch. Structured finance applies securitisation methods and attempts to bring to the financial system the benefits of securitisation such as liquidity and diversification and distribution of risk. The CDO product was one of the products at the centre of the structured finance engine. However, there were failures in governance and risk management. When securitisation is conducted properly in structured finance it can potentially provide advantages to the financial system that outweigh disadvantages. The lessons from the credit crunch should be carefully considered in order to lay firmer foundations for the future. The intention of this chapter is to provide readers with information on the CDO product so that they can engage in this important debate.

Selected bibliography and references

Anson, M. *Credit Derivatives*, FJF Associates, 1999, chapter 3.
Braddock, J. *Derivatives Demystified: Using Structured Finance Products*, Wiley, 1997.
Choudhry, M. 'Some issues in the asset-swap pricing of credit default swaps', *Derivatives Week*, 2 December 2001.
Choudhry, M. 'Trading credit spreads: the case for a specialised exchange-traded credit futures contract', *Journal of Derivatives Use, Trading and Regulation*, 8, 1, June 2002.
Choudhry, M., and G. Landuyt,"Modern portfolio theory and the myth of diversification", *World Commerce Review*, 2009.
Das, S. *Credit Derivatives and Credit-Linked Notes*, 2nd edn, Wiley, 2000.
Das, S. *Structured Products and Hybrid Securities*, Wiley, 2001, chapter 12.
Fabozzi, F. *Handbook of Structured Financial Products*, FJF Associates, 1998.
Fabozzi, F. and Goodman, L. (eds), *Investing in Collateralised Debt Obligations*, FJF Associates, 2001.
ING Barings, *International Asset Securitisation*, 1999.
Kasapi, A. *Mastering Credit Derivatives*, FT Prentice Hall, 1999.
Mattoo, M. *Structured Derivatives*, Pitman, 1996.
Norton Rose, *Collateralised Debt Obligations*, 2000.
Peng, S. and Dattatreya, R. *The Structured Note Market,* Pitman, 1994.

PART IV
Derivative Instruments

Derivative instruments are a large and diverse product class. They are not securities, like the other products discussed in this book (although it is common to see them referred to as such), but are a vital part of the capital markets. The subject matter ordinarily demands several books in its own right. As usual we aim to give readers a flavour of the diversity and introduce the basic techniques of analysis. Hence in Chapter 15 we discuss futures and forward rate agreements. This is followed by a look at swaps. Three chapters are devoted to options, which reflects their complexity. Even so the discussion on options is introductory, and readers are directed to a number of excellent text books in this area. The final chapter looks at credit derivatives.

This page intentionally left blank

15
Short-Term Interest-Rate Derivatives

The market in short-term interest-rate derivatives is large and liquid, and the instruments involved are used for a variety of purposes. Here we review the two main contracts used in money markets trading, the *short-term interest rate future* and the *forward rate agreement*.

Earlier we introduced the concept of the forward rate. Money market derivatives are priced on the basis of the forward rate, and are flexible instruments for hedging against or speculating on forward interest rates. The FRA and the exchange-traded interest-rate future both date from around the same time, and although initially developed to hedge forward interest-rate exposure, they now have a range of uses. In this chapter the instruments are introduced and analysed, and there is a review of the main uses to which they are put. Readers interested in the concept of convexity bias in swap and futures pricing may wish to refer to Choudhry (2004), which is an accessible introduction.

Forward rate agreements

A *forward rate agreement* (FRA) is an over-the-counter (OTC) derivative instrument that trades as part of the money markets. An FRA is essentially a forward-starting loan, dealt at a fixed rate, but with no exchange of cash principal – only the interest applicable on the notional amount between the rate dealt at and the actual rate prevailing at the time of settlement changes hands. In other words, FRAs are *off-balance sheet* (OBS) instruments. By trading today at an interest rate that is effective at some point in the future, FRAs enable banks and corporates to hedge interest rate exposure. They are also used to speculate on the level of future interest rates.

An FRA is an agreement to borrow or lend a *notional* cash sum for a period of time lasting up to 12 months, starting at any point over the next 12 months, at an agreed rate of interest (the FRA rate). The 'buyer' of an FRA is borrowing a notional sum of money while the 'seller' is lending this cash sum. Note how this differs from all other money market instruments. In the cash market, the party buying a collateralised debt instrument or bill, or bidding for stock in the repo market, is the lender of funds. In the FRA market, to 'buy' is to 'borrow'. We use the term

'notional' because with an FRA no borrowing or lending of cash actually takes place, as it is an OBS product. The notional sum is simply the amount on which interest payment is calculated.

So when an FRA is traded, the buyer is borrowing (and the seller is lending) a specified notional sum at a fixed rate of interest for a specified period, the 'loan' to commence at an agreed date in the future. The *buyer* is the notional borrower, and so if there is a rise in interest rates between the date that the FRA is traded and the date that the FRA comes into effect, she will be protected. If there is a fall in interest rates, the buyer must pay the difference between the rate at which the FRA was traded and the actual rate, as a percentage of the notional sum. The buyer may be using the FRA to hedge an actual exposure, that is an actual borrowing of money, or simply speculating on a rise in interest rates. The counterparty to the transaction, the *seller* of the FRA, is the notional lender of funds, and has fixed the rate for lending funds. If there is a fall in interest rates the seller will gain, and if there is a rise in rates the seller will pay. Again, the seller may have an actual loan of cash to hedge or be a speculator.

In FRA trading only the payment that arises as a result of the difference in interest rates changes hands. There is no exchange of cash at the time of the trade. The cash payment that does arise is the difference in interest rates between that at which the FRA was traded and the actual rate prevailing when the FRA matures, as a percentage of the notional amount. FRAs are traded by both banks and corporates and between banks. The FRA market is very liquid in all major currencies, and rates are readily quoted on screens by both banks and brokers. Dealing is over the telephone or over a dealing system such as Reuters.

The terminology quoting FRAs refers to the borrowing time period and the time at which the FRA comes into effect (or matures). Hence if a buyer of a FRA wished to hedge against a rise in rates to cover a three-month loan starting in three months' time, she would transact a 'three-against-six month' FRA, or more usually a 3 × 6 or 3-vs-6 FRA. This is referred to in the market as a 'threes-sixes' FRA, and means a three-month loan in three months' time. So a 'ones-fours' FRA (1 v 4) is a three-month loan in one months' time, and a 'three-nines' FRA (3 v 9) is six-month money in three months' time.

Example 15.1

A company knows that it will need to borrow £1 million in three months' time for a 12-month period. It can borrow funds today at Libor + 50 basis points (bps). Libor rates today are at 5% but the company's treasurer expects rates to go up to about 6% over the next few weeks. So the company will be forced to borrow at higher rates unless some sort of hedge is transacted to protect the borrowing requirement. The treasurer decides to buy a 3 × 15 ('threes-fifteens') FRA to cover the 12 months period beginning three months from now. A bank quotes 5.5% for the FRA, which the company

buys for a notional £1 million. Three months from now rates have indeed gone up to 6%, so the treasurer must borrow funds at 6.5% (the Libor rate plus spread). However she will receive a settlement amount which will be the difference between the rate at which the FRA was bought and today's 12-month Libor rate (6%) as a percentage of £1 million, which will compensate for some of the increased borrowing costs.

In virtually every market FRAs trade under a set of terms and conventions that are identical. The British Bankers Association (BBA) drew up standard legal documentation to cover FRA trading. The following standard terms are used in the market.

NOTIONAL SUM

The amount for which the FRA is traded.

TRADE DATE

The date on which the FRA is dealt.

SETTLEMENT DATE

The date on which the notional loan or deposit of funds becomes effective, that is, is said to begin. This date is used, in conjunction with the notional sum, for calculation purposes only as no actual loan or deposit takes place.

FIXING DATE

This is the date on which the reference rate is determined, that is, the rate to which the FRA dealing rate is compared.

MATURITY DATE

The date on which the notional loan or deposit expires.

CONTRACT PERIOD

The time between the settlement date and maturity date.

FRA RATE

The interest rate at which the FRA is traded.

REFERENCE RATE

This is the rate used as part of the calculation of the settlement amount, usually the Libor rate on the fixing date for the contract period in question.

SETTLEMENT SUM

The amount calculated as the difference between the FRA rate and the reference rate as a percentage of the notional sum, paid by one party to the other on the settlement date.

These terms are illustrated in Figure 15.1.

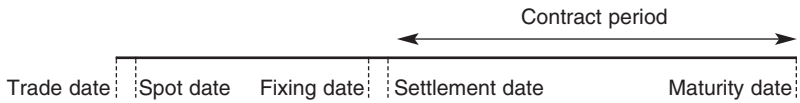


Figure 15.1 Key dates in an FRA trade

The spot date is usually two business days after the trade date; however it can by agreement be sooner or later than this. The settlement date will be the time period after the spot date referred to by the FRA terms, for example a 1 × 4 FRA will have a settlement date one calendar month after the spot date. The fixing date is usually two business days before the settlement date. The settlement sum is paid on the settlement date, and as it refers to an amount over a period of time that is paid up front, at the start of the contract period, the calculated sum is discounted. This is because a normal payment of interest on a loan/deposit is paid at the end of the time period to which it relates; because an FRA makes this payment at the start of the relevant period, the settlement amount is a discounted figure.

With most FRA trades the reference rate is the Libor fixing on the fixing date.

The settlement sum is calculated after the fixing date, for payment on the settlement date. We may illustrate this with an hypothetical example. Consider a case where a corporate has bought £1 million notional of a 1v4 FRA, and dealt at 5.75%, and that the market rate is 6.50% on the fixing date. The contract period is 90 days. In the cash market the extra interest charge that the corporate would pay is a simple interest calculation, and is:

$$\frac{6.50-5.75}{100}\times 1{,}000{,}000\times\frac{91}{365}=\text{£}1869.86$$

This extra interest that the corporate is facing would be payable with the interest payment for the loan, which (as it is a money market loan) is when the loan matures. Under a FRA, then, the settlement sum payable should, if it was paid on the same day as the cash market interest charge, be exactly equal to this. This would make it a perfect hedge. As we noted above, though, FRA settlement value is paid at the start of the contract period, that is, the beginning of the underlying loan and not the end. Therefore the settlement sum has to be adjusted to account for this, and the amount of the adjustment is the value of the interest that would be earned if the unadjusted cash value was invested for the contract period in the money market. The amount of the settlement value is given by (15.1).

$$Settlement=\frac{(r_{ref}-r_{FRA})\times M\times\frac{n}{B})}{1+\left(r_{ref}\times\frac{n}{B}\right)} \tag{15.1}$$

where

r_{ref} is the reference interest fixing rate
r_{FRA} is the FRA rate or *contract rate*
M is the notional value
n is the number of days in the contract period
B is the day-count base (360 or 365)

The expression at (15.1) simply calculates the extra interest payable in the cash market, resulting from the difference between the two interest rates, and then discounts the amount because it is payable at the start of the period and not, as would happen in the cash market, at the end of the period.

In our hypothetical illustration, as the fixing rate is higher than the dealt rate, the corporate buyer of the FRA receives the settlement sum from the seller. This then compensates the corporate for the higher borrowing costs that he would have to pay in the cash market. If the fixing rate had been lower than 5.75%, the buyer would pay the difference to the seller, because the cash market rates mean that he is subject to a lower interest rate in the cash market. What the FRA has done is hedge the corporate, so that whatever happens in the market, it will pay 5.75% on its borrowing.

A market maker in FRAs is trading short-term interest rates. The settlement sum is the value of the FRA. The concept is exactly as with trading short-term interest-rate futures: a trader who buys a FRA is running a long position, so that if on the fixing date $r_{ref} > r_{FRA}$, the settlement sum is positive and the trader realises a profit. What has happened is that the trader, by buying the FRA, 'borrowed' money at an interest rate, which subsequently rose. This is a gain, exactly like a *short* position in an interest-rate future, where if the price goes down (that is, interest rates go up), the trader realises a gain. Equally a 'short' position in a FRA, put on by selling a FRA, realises a gain if on the fixing date $r_{ref} < r_{FRA}$.

FRA pricing

As their name implies, FRAs are forward rate instruments and are priced using the forward rate principles we established earlier in the book.

We use the standard forward-rate breakeven formula to solve for the required FRA rate. The relationship given at (15.2) connects simple (bullet) interest rates for periods of time up to one year, where no compounding of interest is required. As FRAs are money market instruments we are not required to calculate rates for periods in excess of one year,[1] where compounding would need to built into the equation.

$$(1 + r_2t_2) = (1 + r_1t_1)(1 + r_ft_f) \tag{15.2}$$

[1]It is of course possible to trade FRAs with contract periods greater than one year, for which a different pricing formula must be used.

where

r_2 is the cash market interest rate for the long period
r_1 is the cash market interest rate for the short period
r_f is the forward rate for the gap period
t_2 is the time period from today to the end of the long period
t_1 is the time period from today to the end of the short period
t_f is the forward gap time period, or the contract period for the FRA.

This is illustrated diagrammatically in Figure 15.2.

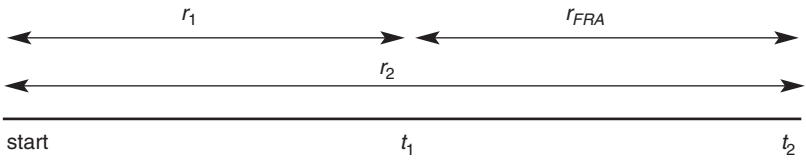


Figure 15.2 Rates used in FRA pricing

The time period t_1 is the time from the dealing date to the FRA settlement date, while t_2 is the time from the dealing date to the FRA maturity date. The time period for the FRA (contract period) is t_2 minus t_1. We can replace the symbol t for time period with n for the actual number of days in the time periods themselves. If we do this and then rearrange the equation to solve for r_{FRA}, the FRA rate, we obtain (15.3):

$$r_{FRA} = \frac{r_2 n_2 - r_1 n_1}{n_{FRA}\left(1 + r_1 \dfrac{n_1}{365}\right)} \tag{15.3}$$

where

n_1 is the number of days from the dealing date or spot date to the settlement date
n_2 is the number of days from dealing date or spot date to the maturity date
r_1 is the spot rate to the settlement date
r_2 is the spot rate from the spot date to the maturity date
n_{FRA} is the number of days in the FRA contract period
r_{FRA} is the FRA rate

If the formula is applied to say the US money markets, the 365 in the equation is replaced by 360, the day count base for that market.

In practice FRAs are priced off the exchange-traded short-term interest rate future for that currency, so that sterling FRAs are priced off LIFFE short sterling futures. Traders normally use a spreadsheet pricing model that has futures prices

directly fed into it. FRA positions are also usually hedged with other FRAs or short-term interest rate futures.

FRA prices in practice

The dealing rates for FRAs are possibly the most liquid and transparent of any non-exchange traded derivative instrument. This is because they are calculated directly from exchange-traded interest-rate contracts. The key consideration for FRA market makers however is how the rates behave in relation to other market interest rates. The forward rate calculated from two period spot rates must, as we have seen, be set such that it is arbitrage-free. If for example the six-month spot rate was 8.00% and the nine-month spot rate was 9.00%, the 6 v 9 FRA would have an approximate rate of 11%. What would be the effect of a change in one or both of the spot rates? The same arbitrage-free principle must apply. If there is an increase in the short-rate period, the FRA rate must decrease, to make the total return unchanged. The extent of the change in the FRA rate is a function of the ratio of the contract period to the long period. If the rate for the long period increases, the FRA rate will increase, by an amount related to the ratio between the total period to the contract period. The FRA rate for any term is generally a function of the three-month Libor rate, usually the rate traded under an interest-rate future. A general rise in this rate will see a rise in FRA rates.

Forward contracts

A forward contract is an OTC instrument with terms set for delivery of an underlying asset at some point in the future. That is, a forward contract fixes the price and the conditions now for an asset that will be delivered in the future. As each contract is tailor-made to suit user requirements, a forward contract is not as liquid as an exchange-traded futures contract with standardised terms.

The theoretical textbook price of a forward contract is the spot price of the underlying asset plus the funding cost associated with holding the asset until forward expiry date, when the asset is delivered. More formally it can be shown[2] that the price of a forward contract (written on an underlying asset that pays no dividends, such as a zero-coupon bond), is as (15.4).

$$P_{fwd} = P_{und}e^{rn} \tag{15.4}$$

where

P_{und} is the price of the underlying asset of the forward contract
r is the continuously compounded risk-free interest rate for a period of maturity n
n is the term to maturity of the forward contract in days

[2] For instance see Hull (1999), Jarrow and Turnbull (2000) or Kolb (2000).

The rule of no-arbitrage pricing states that (15.4) must be true. If $P_{fwd} < P_{und}e^{rn}$ then a trader could buy the cheaper instrument, the forward contract, and simultaneously sell the underlying asset. The proceeds from the short sale could be invested at r for n days; on expiry the short position in the asset is closed out at the forward price P_{fwd} and the trader will have generated a profit of $P_{und}e^{rn} - P_{fwd}$. In the opposite scenario, where $P_{fwd} > P_{und}e^{rn}$, a trader could put on a long position in the underlying asset, funded at the risk-free interest rate r for n days, and simultaneously sell the forward contract. On expiry the asset is sold under the terms of the forward contract at the forward price and the proceeds from the sale are used to close out the funding initially taken on to buy the asset. Again a profit would be generated, which would be equal to the difference between the two prices.

The relationship described here is used by the market to assume that forward rates implied by the price of short-term interest-rate futures contracts are equal to forward rates given by a same-maturity forward contract. Although this assumption holds good for futures contracts with a maturity of up to three or four years, it breaks down for longer-dated futures and forwards. An accessible account of this feature is contained in Hull (1999).

Short-term interest rate futures

Description

A *futures* contract is a transaction that fixes the price today for a commodity that will be delivered at some point in the future. Financial futures fix the price for interest rates, bonds, equities and so on, but trade in the same manner as commodity futures. Contracts for futures are standardised and traded on exchanges. In London the main futures exchange is LIFFE, although commodity futures are also traded on for example, the International Petroleum Exchange and the London Metal Exchange. The money markets trade short-term interest rate futures, which fix the rate of interest on a notional fixed term deposit of money (usually for 90 days or three months) for a specified period in the future. The sum is notional because no actual sum of money is deposited when buying or selling futures; the instrument is off-balance sheet. Buying such a contract is equivalent to making a notional deposit, while selling a contract is equivalent to borrowing a notional sum.

The three-month interest-rate future is the most widely used instrument used for hedging interest-rate risk.

The LIFFE exchange in London trades short-term interest rate futures for major currencies including sterling, euros, yen and Swiss francs. Table 15.1 summarises the terms for the short sterling contract as traded on LIFFE.

The original futures contracts related to physical commodities, which is why we speak of *delivery* when referring to the expiry of financial futures contracts. Exchange-traded futures such as those on LIFFE are set to expire every quarter during the year. The short sterling contract is a deposit of cash, so as its price refers to the rate of interest on this deposit, the price of the contract is set as $P = 100 - r$, where P is the price of the contract and r is the rate of interest at the time of expiry implied by the futures contract. This means that if the price of the contract rises,

Table 15.1 Description of LIFFE short sterling future contract

Name	*90-day sterling Libor future*
Contract size	£500,000
Delivery months	March, June, September, December
Delivery date	First business day after last trading day
Last trading day	Third Wednesday of delivery month
Price	100 minus yield
Tick size	0.01
Tick value	£12.50
Trading hours	0805–1757 (Electronic screen trading)

the rate of interest implied goes down, and vice versa. For example the price of the June 2009 short sterling future (written as Jun09 or M09, from the futures identity letters of H, M, U and Z for contracts expiring in March, June, September and December respectively) at the start of trading on 17 March 2009 was 98.35, which implied a three-month Libor rate of 1.65% on expiry of the contract in June. If a trader bought 20 contracts at this price and then sold them just before the close of trading that day, when the price had risen to 98.84, an implied rate of 1.16%, she would have made 49 ticks profit or £12,250. That is, a 20 tick upward price movement in a long position of 20 contracts is equal to £12,250. This is calculated as follows:

$$\text{Profit} = \text{ticks gained} \times \text{tick value} \times \text{number of contracts}$$
$$\text{Loss} = \text{ticks lost} \times \text{tick value} \times \text{number of contracts}$$

The tick value for the short sterling contract is straightforward to calculate, since we know that the contract size is £500,000, there is a minimum price movement (tick movement) of 0.01% and the contract has a three month 'maturity'.

$$\text{Tick value} = 0.01\% \times £500{,}000 \times (3/12) = £12.50$$

The profit made by the trader in our example is logical because if we buy short sterling futures we are depositing (notional) funds; if the price of the futures rises, it means the interest rate has fallen. We profit because we have 'deposited' funds at a higher rate beforehand. If we expected sterling interest rates to rise, we would sell short sterling futures, which is equivalent to borrowing funds and locking in the loan rate at a lower level.

Note how the concept of buying and selling interest rate futures differs from FRAs: if we buy an FRA we are borrowing notional funds, whereas if we buy a futures contract we are depositing notional funds. If a position in an interest rate futures contract is held to expiry, cash settlement will take place on the delivery day for that contract.

Short-term interest rate contracts in other currencies are similar to the short sterling contract, and trade on exchanges such as Deutsche Terminbourse in Frankfurt and MATIF in Paris.

Pricing interest rate futures

The price of a three-month interest rate futures contract is the implied interest rate for that currency's three-month rate at the time of expiry of the contract. Therefore there is always a close relationship and correlation between futures prices, FRA rates (which are derived from futures prices) and cash market rates. On the day of expiry, the price of the future will be equal to the Libor rate as fixed that day. This is known as the *exchange delivery settlement price* (EDSP), and is used in the calculation of the delivery amount. During the life of the contract its price will be less closely related to the actual three-month Libor rate *today*, but closely related to the *forward rate* for the time of expiry.

Equation (15.2) was our basic forward rate formula for money market maturity forward rates, which we adapted to use as our FRA price equation. If we incorporate some extra terminology to cover the dealing dates involved it can also be used as our futures price formula. Assume that:

T_0 is the trade date
T_M is the contract expiry date
T_{CASH} is the value date for cash market deposits traded on T_0
T_1 is the value date for cash market deposits traded on T_M
T_2 is the maturity date for a three-month cash market deposit traded on T_M.

We can then use equation (15.2) as our futures price formula to obtain P_{fut}, the futures price for a contract up to the expiry date.

$$P_{fut} = 100 - \left[\frac{r_2 n_2 - r_1 n_1}{n_f \left(1 + r_1 \frac{n_1}{365} \right)} \right] \tag{15.5}$$

where

P_{fut} is the futures price
r_1 is the cash market interest rate to T_1
r_2 is the cash market interest rate to T_2
n_1 is the number of days from T_{CASH} to T_1
n_2 is the number of days from T_{CASH} to T_2
n_f is the number of days from T_1 to T_2

The formula uses a 365-day count convention which applies in the sterling money markets; where a market uses a 360-day base this must be used in the equation instead.

In practice the price of a contract at any one time will be close to the theoretical price that would be established by (15.5). Discrepancies will arise for supply and demand reasons in the market, as well as because Libor rates are often quoted only to the nearest sixteenth or 0.0625. The prices of FRAs and futures are correlated

very closely, in fact banks will often price FRAs using futures, and use futures to hedge their FRA books. When hedging a FRA book with futures, the hedge is quite close to being exact, because the two prices track each other almost tick for tick. However the tick value of a futures contract is fixed, and uses (as we saw above) a 3/12 basis, while FRA settlement values use a 360 or 365 day base. The FRA trader will be aware of this when putting on her hedge.

In the discussion on forward rates we emphasised that they were the market's view on future rates using all information available today. Of course a futures price today is very unlikely to be in line with the actual three-month interest rate that is prevailing at the time of the contract's expiry. This explains why prices for futures and actual cash rates will differ on any particular day. Up until expiry the futures price is the implied forward rate; of course there is always a discrepancy between this forward rate and the cash market rate *today*. The gap between the cash price and the futures price is known as the basis. This is defined as:

$$\text{Basis} = \text{Cash price} - \text{Futures price}$$

At any point during the life of a futures contract prior to final settlement – at which point futures and cash rates converge – there is usually a difference between current cash market rates and the rates implied by the futures price. This is the difference we have just explained. In fact the difference between the price implied by the current three-month interbank deposit and the futures price is known as *simple basis*, but it is what most market participants refer to as the basis. Simple basis consists of two separate components, *theoretical basis* and *value basis*. Theoretical basis is the difference between the price implied by the current three-month interbank deposit rate and that implied by the theoretical fair futures price based on cash market forward rates, given by (15.5). This basis may be either positive or negative depending on the shape of the yield curve.

The value basis is the difference between the theoretical fair futures price and the actual futures price. It is a measure of how under or over-valued the futures contract is relative to its fair value. Value basis reflects the fact that a futures contract does not always trade at its mathematically calculated theoretical price, due to the impact of market sentiment and demand and supply. The theoretical and value can and do move independently of one another and in response to different influences. Both however converge to zero on the last trading day when final cash settlement of the futures contract is made.

Futures contracts do not in practice provide a precise tool for locking into cash market rates today for a transaction that takes place in the future, although this is what they are in theory designed to do. Futures do allow a bank to lock in a rate for a transaction to take place in the future, and this rate is the *forward rate*. The basis is the difference between today's cash market rate and the forward rate on a particular date in the future. As a futures contract approaches expiry, its price and the rate in the cash market will converge (the process is given the name *convergence*). As we noted earlier this is given by the EDSP, and the two prices (rates) will be exactly in line at the exact moment of expiry.

Hedging using interest-rate futures

Banks use interest rate futures to hedge interest rate risk exposure in cash and OBS instruments. Bond trading desks also often use futures to hedge positions in bonds of up to two or three years maturity, as contracts are traded up to three years maturity. The liquidity of such 'far month' contracts is considerably lower than for near month contracts and the 'front month' contract (the current contract, for the next maturity month). When hedging a bond with a maturity of say two years maturity, the trader will put on a *strip* of futures contracts that matches as near as possible the expiry date of the bond.

The purpose of a hedge is to protect the value of a current or anticipated cash market or OBS position from adverse changes in interest rates. The hedger will try to offset the effect of the change in interest rate on the value of his cash position with the change in value of her hedging instrument. If the hedge is an exact one, the loss on the main position should be compensated by a profit on the hedge position. If the trader is expecting a fall in interest rates and wishes to protect against such a fall, she will buy futures, known as a long hedge, and will sell futures (a short hedge) if wishing to protect against a rise in rates.

Bond traders also use three-month interest rate contracts to hedge positions in short-dated bonds; for instance, a market maker running a short-dated bond book would find it more appropriate to hedge his book using short-dated futures rather than the longer-dated bond futures contract. When this happens it is important to accurately calculate the correct number of contracts to use for the hedge. To construct a bond hedge it will be necessary to use a strip of contracts, thus ensuring that the maturity date of the bond is covered by the longest-dated futures contract. The hedge is calculated by finding the sensitivity of each cash flow to changes in each of the relevant forward rates. Each cash flow is considered individually, and the hedge values are then aggregated and rounded to the nearest whole number of contracts.

The following examples illustrate hedging with short-term interest-rate contracts.

Example 15.2: Hedging a forward three-month lending requirement

On 1 June a corporate treasurer is expecting a cash inflow of £10 million in three months time (1 September), which she will then invest for three months. The treasurer expects that interest rates will fall over the next few weeks and wishes to protect herself against such a fall. This can be done using short sterling futures. Market rates on 1 June are as follows:

3-month Libor	6.5%
Sept. futures price	93.220

The treasurer buys 20 September short sterling futures at 93.220, this number being exactly equivalent to a sum of £10 million. This allows her to lock

in a forward lending rate of 6.78%, if we assume there is no bid–offer quote spread.

Expected lending rate = rate implied by futures price
= 100 − 93.220 = 6.78%

On 1 September market rates are as follows:

3-month Libor	6¼%
Sept. futures price	93.705

The treasurer unwinds the hedge at this price.

Futures p/l = + 97 ticks (93.705 − 93.22), or 0.485%

Effective lending rate = 3-month Libor + futures profit
= 6.25% + 0.485% = 6.735%

The treasurer was quite close to achieving her target lending rate of 6.78%, and the hedge has helped to protect against the drop in Libor rates from 6½% to 6¼%, due to the profit from the futures transaction.

In the real world the cash market bid–offer spread will impact the amount of profit/loss from the hedge transaction. Futures generally trade and settle near the offered side of the market rate (Libor), whereas lending, certainly by corporates, will be nearer the Libid rate.

Example 15.3: Hedging a forward six-month borrowing requirement

A treasury dealer has a six-month borrowing requirement for €30 million in three months time, on 16 September. She expects interest rates to rise by at least 0.5% before that date and would like to lock in a future borrowing rate. The scenario is detailed below.

Date:	16 June
Three-month Libor	6.0625%
Six-month Libor	6.25
Sept. futures contract	93.66
Dec. futures contract	93.39

In order to hedge a six-month €30 million exposure the dealer needs to use a total of 60 futures contracts, as each has a nominal value of €1 million, and corresponds to a three-month notional deposit period. The dealer decides to sell 30 September futures contracts and 30 December futures contracts,

which is referred to as a *strip hedge*. The expected forward borrowing rate that can be achieved by this strategy, where the expected borrowing rate is *rf*, is calculated as follows:

$$\left[1+rf\times\frac{\text{days in period}}{360}\right]=\left[1+\text{Sept. implied rate}\times\frac{\text{Sept. days period}}{360}\right]$$

$$\times\left[1+\text{Dec. implied rate}\times\frac{\text{Dec. days period}}{360}\right]$$

Therefore we have:

$$\left[1+rf\times\frac{180}{360}\right]=\left[1+0.0634\times\frac{90}{360}\right]\times\left[1+0.0661\times\frac{90}{360}\right]$$

The rate *rf* is sometimes referred to as the 'strip rate'.

The hedge is unwound upon expiry of the September futures contract. Assume the following rates now prevail:

Three-month Libor	6.4375%
Six-month Libor	6.8125
Sept. futures contract	93.56
Dec. futures contract	92.93

The futures profit and loss is:

September contract:	+ 10 ticks
December contract:	+ 46 ticks

This represents a 56 tick or 0.56% profit in three-month interest-rate terms, or 0.28% in six-month interest-rate terms. The effective borrowing rate is the six-month Libor rate minus the futures profit, 6.8125% – 0.28% or 6.5325%.

In this case the hedge has proved effective because the dealer has realised a borrowing rate of 6.5325%, which is close to the target strip rate of 6.53%.

The dealer is still exposed to the basis risk when the December contracts are bought back from the market at the expiry date of the September contract. If, for example, the future was bought back at 92.73, the effective borrowing rate would be only 6.4325%, and the dealer would benefit. Of course the other possibility is that the futures contract could be trading 20 ticks more expensive, which would give a borrowing rate of 6.6325%, which is 10 basis points above the target rate. If this happened, the dealer may elect to borrow in the cash market for three months, and maintain the December futures position until the December contract expiry date, and roll over the borrowing at that time. The profit (or loss) on the December futures position will compensate for any change in three-month rates at that time.

Refining the hedge ratio

A futures hedge ratio is calculated by dividing the amount to be hedged by the nominal value of the relevant futures contract and then adjusting for the duration of the hedge. When dealing with large exposures and/or a long hedge period, inaccuracy will result unless the hedge ratio is refined to compensate for the timing mismatch between the cash flows from the futures hedge and the underlying exposure. Any change in interest rates has an immediate effect on the hedge in the form of daily variation margin, but only affects the underlying cash position on maturity, that is, when the interest payment is due on the loan or deposit. In other words, hedging gains and losses in the futures position are realised over the hedge period while cash market gains and losses are deferred. Futures gains may be reinvested, and futures losses need to financed.

The basic hedge ratio is usually refined to counteract this timing mismatch, this process is sometimes called 'tailing'.

Example 15.4: Refining the hedge ratio

A dealer is hedging a three-month SFr100 million borrowing commitment commencing in two months' time and wishes to determine an accurate hedge ratio.

Two-month Libor	4.75%
Five-month Libor	4.875%
Implied 2 v 5 fwd-fwd rate	4.92%

The first part of the process to refine the hedge ratio involves measuring the sensitivity of the underlying position to a change in interest rates, that is, the cost of a basis point move in Libor on the interest payment or receipt on maturity. We therefore calculate the basis point value as follows:

$$\text{Principal} \times 0.01\% \times \frac{\text{term (days)}}{360}$$

In this case the PVBP is SFr2,500.

For every basis point increase (decrease) in three-month Libor, the dealer's interest-rate expense at the maturity of the loan will increase (decrease) by SFr2500. Correspondingly a SFr2,500 gain (loss) will be realised over the hedge period on the futures position. The present value of the futures gain (loss) is therefore greater than the present value of the loss (gain) on the loan. In other words, a hedge position consisting of short 100 lots is an over-hedge.

To calculate a more precise hedge ratio, the dealer needs to discount the nominal basis point value of the interest payments back from the maturity date of the loan to the start date of the loan. The discounting rate used in this calculation is the forward-forward rate over the loan period, that is, three months, implied by the current Libor market rates (given here as the

2 v 5 fwd-fwd rate). The discounting period may vary depending on assumptions about the timing of cash flows. The formula for calculating the present value at the start date of the loan of a basis point move is calculated as:

$$\frac{\text{Nominal value of a basis point}}{1+\text{fwd}-\text{fwd rate}\times\dfrac{\text{loan period (days)}}{360}}$$

For the dealer the calculation is:

$$\frac{250}{\left[1+4.92\%\times\dfrac{90}{360}\right]}=\text{SFr}2469.6$$

To obtain the hedge ratio from this figure, the dealer would divide the PVBP value by the tick value of the futures contract, which for the LIFFE Euroswiss contract is SFr25:

$$\frac{\text{SFr}2,470}{\text{SFr}25}=98.80$$

Therefore the correct number of contracts needed to hedge the SFr100 million exposure would be 99, rather than 100.

Appendix 15.1: The forward interest rate and futures-implied forward rate

The markets assume that the forward rate implied by the price of a futures contract is the same as the futures price itself for a contract with the same expiry date. This assumption is the basis on which futures contracts are used to price swaps and other forward rate instruments such as FRAs. In Appendix 15.2 we summarise a strategy first described by Cox, Ingersoll and Ross (1981) to show that under certain assumptions, namely when the risk-free interest rate is constant and identical for all maturities (that is, in a flat term structure environment), this assumption holds true. However in practice, because the assumptions are not realistic under actual market conditions, this relationship does not hold for longer-dated futures contracts and forward rates.

In the first place, term structures are rarely flat or constant. The main reason however is because of the way futures contracts are settled, compared with forward contracts. Market participants who deal in exchange-traded futures must deposit daily margin with the exchange clearing house, reflecting their profit and loss on futures trading. Therefore a profit on a futures position will be received immediately, and in a positive-sloping yield curve environment this profit will be invested at a higher-than-average rate of interest. In the same way a loss on futures trading would have to be funded straight away, and the funding cost would be at a lower-than-average rate of interest. However the profit on a forward contract is not realised until the maturity of the contract, and so a position in a forward is not affected by daily profit or loss cash flows. Therefore, a long-dated futures contract will have more value to

an investor than a long-dated forward contract, because of the opportunity to invest mark-to-market gains made during the life of the futures contract.

When the price of the underlying asset represented by a futures contract is positively correlated with interest rates, the price of futures contracts will be higher than the price of the same-contract forward contract. When the price of the underlying asset is negatively correlated with interest rates, which is the case with three-month interest-rate futures like short sterling, forward prices are higher than futures prices. That is, the forward interest rate is lower than the interest rate implied by the futures contract price. This difference is not pronounced for short-dated contracts, and so is ignored by the market.

There are also other factors that will cause a difference in forward and futures prices, the most significant of these being transaction costs and liquidity: it is generally cheaper to trade exchange-traded futures and they tend to be more liquid instruments. However for longer-dated instruments, the difference in treatment between forwards and futures means that their rates will not be the same, and this difference needs to be taken into account when pricing long-dated forward instruments. The issue of convexity bias is discussed in a number of texts, the best-known being Hull (1999).

Appendix 15.2: Arbitrage proof of the futures price being equal to the forward price

Under certain assumptions it can be shown that the price of same-maturity futures and forward contracts are equal. The primary assumption is that interest rates are constant. The strategy used to prove this was first described by Cox, Ingersoll and Ross (1981).

Consider a futures contract with maturity of n days and with a price of P_i at the end of day i. Set r as the constant risk-free interest rate of interest per day. Assume a trading strategy that consists of:

- establishing a long position in the futures of e^r at the start of day 0
- adding to the long position to make a total of e^{2r} at the end of day 1
- adding to the long position to make a total of e^{3r} at the end of day 2
- increasing the size of the position daily by the amount shown.

At the start of day i the long position is e^{ir}. The profit or loss from the position is given by (15.6).

$$P/L = (P_i - P_{i-1})e^{ir} \tag{15.6}$$

If this amount is compounded on a daily basis using r, the final value on the expiry of the contract is given by:

$$(P_i - P_{i-1})e^{ir}\, e^{(n-i)r} = (P_i - P_{i-1})e^{nr}$$

so that the value of the position on the expiry of the contract at the end of day n is given by (15.7)

$$FV = \sum_{i=1}^{n} (P_i - P_{i-1}) e^{nr} \tag{15.7}$$

The expression at (15.7) may also be written as (15.8):

$$FV = [(P_n - P_{n-1}) + (P_{n-1} - P_{n-2}) + ... + (P_1 - P_0)]\, e^{nr} = (P_n - P_0)\, e^{nr} \tag{15.8}$$

In theory the price of a futures contract on expiry must equal the price of the underlying asset on that day. If we set the price of the underlying asset on expiry as $P_{n\text{-}underlying,}$ as P_n is equal to the final price of the contract on expiry, the final value of the trading strategy may be written as (15.9):

$$FV = (P_{n\text{–}underlying} - P_0)\, e^{nr} \tag{15.9}$$

Investing P_0 in a risk-free bond and using the same strategy as that described above will therefore return:

$$P_0\, e^{nr} + (P_{n\text{–}underlying} - P_0)\, e^{nr}$$

or an amount equal to $P_{n\text{-}underlying}e^{nr}$ at the expiry of the contract at the close of day n. Therefore this states than an amount P_0 may be invested to return a final amount of $P_{n\text{-}underlying}e^{nr}$ at the end of day n.

Assume that the forward contract price at the end of day 0 is $P_{0\text{-}forward}$. By investing this amount in a risk-free bond, and simultaneously establishing a long forward position of e^{nr} forward contracts, we are guaranteed an amount $P_{n\text{-}underlying}e^{nr}$ at the end of day n. We therefore have two investment strategies that both return a value of $P_{n\text{-}underlying}e^{nr}$ at the end of the same time period; one strategy requires an investment of P_0 while the other requires an investment of $P_{0\text{-}forward}$. Under the rule of no-arbitrage pricing, the price of both contracts must be equal, that is, $P_0 = P_{0\text{-}forward}$. That is, the price of the futures contract and the price of the forward contract at the end of day 0 are equal.

Selected bibliography and references

Chicago Board of Trade, *Interest rate Futures for Institutional Investors*, CBOT, 1987.

Choudhry, M. *Fixed Income Markets*, John Wiley Asia, 2005, ch. 15.

Cox, J., Ingersoll, J. and Ross, S. 'The relationship between forward prices and futures prices', *Journal of Financial Economics*, 9, Dec. 1981, pp. 321–46.

Figlewski, F. *Hedging with Financial Futures for Institutional Investors*, Probus Publishing, 1986.

French, K. 'A comparison of futures and forwards prices', *Journal of Financial Economics*, 12, Nov. 1983, pp. 311–42.

Hull, J. *Options, Futures and Other Derivatives*, 4th edition, Prentice-Hall, 1999.

Jarrow, R. and Oldfield, G. 'Forward contracts and futures contracts', *Journal of Financial Economics*, 9, Dec. 1981, pp. 373–82.

Jarrow, R. and Turnbull, S. *Derivative Securities*, 2nd edn, South-Western, 2000.

Kolb, R. *Futures, Options and Swaps*, 3rd edn, Blackwell, 2000.

Veit, W. and Reiff, W. 'Commercial banks and interest rate futures: a hedging survey', *Journal of Futures Markets*, 3, 1983, pp. 283–93.

16
Swaps

Swaps are one of the most important and useful instruments in the debt capital markets. They are used by a wide range of institutions, including banks, mortgage banks and building societies, corporates and local authorities. The demand for them has grown as the continuing uncertainty and volatility of interest rates and exchange rates has made it more important to hedge exposures. As the market has matured the instrument has gained wider acceptance, and is regarded as a 'plain vanilla' product in the debt capital markets.

Virtually all commercial and investment banks will quote swap prices for their customers, and as they are over-the-counter (OTC) instruments, dealt over the telephone, it is possible for banks to tailor swaps to match the precise requirements of individual customers. There is also a close relationship between the bond market and the swap market, and corporate finance teams and underwriting banks keep a close eye on the government yield curve and the swap yield curve, looking out for interest-rate advantages and other possibilities regarding new issue of debt.

We do not propose to cover the historical evolution of the swaps markets, which is abundantly covered in existing literature, or the myriad of swap products which can be traded today (ditto). Instead we review the use of interest-rate swaps from the point of view of the bond market participant; this includes pricing and valuation and its use as a hedging tool. There is also an introduction to currency swaps and swaptions. The bibliography lists further reading on important topics such as pricing, valuation and credit risk.

Interest rate swaps

Introduction

Interest-rate swaps are the most important type of swap in terms of volume of transactions. They are used to manage and hedge interest rate risk and exposure, while market makers will also take positions in swaps that reflect their view on the direction of interest rates. An interest rate swap is an agreement between two

counterparties to make periodic interest payments to one another during the life of the swap, on a pre-determined set of dates, based on a *notional* principal amount. One party is the fixed-rate payer, and this rate is agreed at the time of trade of the swap; the other party is the floating-rate payer, the floating rate being determined during the life of the swap by reference to a specific market index.

The principal or notional amount is never physically exchanged, hence the term 'off-balance sheet', but is used merely to calculate the interest payments. The fixed-rate payer receives floating-rate interest and is said to be 'long' or to have 'bought' the swap. The long side has conceptually purchased a floating-rate note (because it receives floating-rate interest) and issued a fixed coupon bond (because it pays out fixed interest at intervals), that is, it has in principle borrowed funds. The floating-rate payer is said to be 'short' or to have 'sold' the swap. The short side has conceptually purchased a coupon bond (because it receives fixed-rate interest) and issued a floating-rate note (because it pays floating-rate interest). So an interest rate swap is:

- an agreement between two parties
- to exchange a stream of cash flows
- calculated as a percentage of a *notional* sum
- and calculated on different interest bases.

For example in a trade between Bank A and Bank B, Bank A may agree to pay fixed semi-annual coupons of 10% on a notional principal sum of £1 million, in return for receiving from Bank B the prevailing six-month sterling Libor rate on the same amount. The known cash flow is the fixed payment of £50,000 every six months by Bank A to Bank B.

Interest-rate swaps trade in a secondary market so their value moves in line with market interest rates, in exactly the same way as bonds. If a five-year interest-rate swap is transacted today at a rate of 5%, and five-year interest rates subsequently fall to 4.75%, the swap will have decreased in value to the fixed-rate payer, and correspondingly increased in value to the floating-rate payer, who has now seen the level of interest payments fall. The opposite would be true if five-year rates moved to 5.25%. Why is this? Consider the fixed-rate payer in an IR swap to be a borrower of funds; if she fixes the interest rate payable on a loan for five years, and then this interest rate decreases shortly afterwards, is she better off? No, because she is now paying above the market rate for the funds borrowed. For this reason a swap contract decreases in value to the fixed-rate payer if there is a fall in rates. Equally a floating-rate payer gains if there is a fall in rates, as he can take advantage of the new rates and pay a lower level of interest; hence the value of a swap increases to the floating-rate payer if there is a fall in rates.

A bank swaps desk will have an overall net interest rate position arising from all the swaps it has traded that are currently on the book. This position is an interest rate exposure at all points along the term structure, out to the maturity of the

longest-dated swap. At the close of business each day all the swaps on the book will be *marked-to-market* at the interest rate quote for that day.

A swap can be viewed in two ways, either as a bundle of forward or futures contracts, or as a bundle of cash flows arising from the 'sale' and 'purchase' of cash market instruments. If we imagine a strip of futures contracts, maturing every three or six months out to three years, we can see how this is conceptually similar to a three-year interest-rate swap. However in our view it is better to visualise a swap as being a bundle of cash flows arising from cash instruments.

Let us imagine we have only two positions on our book:

- a long position in £100 million of a three-year floating-rate note (FRN) that pays six-month Libor semi-annually, and is trading at par
- a short position in £100 million of a three-year gilt with coupon of 6% that is also trading at par.

Being short a bond is the equivalent to being a borrower of funds. Assuming this position is kept to maturity, the resulting cash flows are shown in Table 16.1.

There is no net outflow or inflow at the start of these trades, as the £100 million purchase of the FRN is netted with receipt of £100 million from the sale of the gilt. The resulting cash flows over the three-year period are shown in the last column of Table 16.1. This net position is exactly the same as that of a fixed-rate payer in an interest rate (IR) swap. As we had at the start of the trade, there is no cash inflow or outflow on maturity. For a floating-rate payer, the cash flow would mirror exactly a long position in a fixed-rate bond and a short position in an FRN. Therefore the fixed-rate payer in a swap is said to be short in the bond market, that is a borrower of funds; the floating-rate payer in a swap is said to be long the bond market.

Market terminology

Virtually all swaps are traded under the legal terms and conditions stipulated in the ISDA standard documentation. The trade date for a swap is, not surprisingly,

Table 16.1 Three-year cash flows

Cash flows resulting from long position in FRN and short position in gilt

Period (6 mo)	*FRN*	*Gilt*	*Net cash flow*
0	−£100m	+£100m	£0
1	+(Libor × 100)/2	−3	+(Libor × 100)/2 − 3.0
2	+(Libor × 100)/2	−3	+(Libor × 100)/2 − 3.0
3	+(Libor × 100)/2	−3	+(Libor × 100)/2 − 3.0
4	+(Libor × 100)/2	−3	+(Libor × 100)/2 − 3.0
5	+(Libor × 100)/2	−3	+(Libor × 100)/2 − 3.0
6	+[(Libor × 100)/2] + 100	−103	+(Libor × 100)/2 − 3.0

The Libor rate is the six–month rate prevailing at the time of the setting, for instance the Libor rate at period 4 will be the rate actually prevailing at period 4

the date on which the swap is transacted. The terms of the trade include the fixed interest rate, the maturity and notional amount of the swap, and the payment bases of both legs of the swap. The date from which floating interest payments are determined is the *setting date*, which may also be the trade date. Most swaps fix the floating-rate payments to Libor, although other reference rates that are used include the US Prime rate, Euribor, the Treasury bill rate and the commercial paper rate. In the same way as for FRA and eurocurrency deposits, the rate is fixed two business days before the interest period begins. The second (and subsequent) setting date will be two business days before the beginning of the second (and subsequent) swap periods. The effective date is the date from which interest on the swap is calculated, and this is typically two business days after the trade date. In a forward-start swap the effective date will be at some point in the future, specified in the swap terms. The floating interest-rate for each period is fixed at the start of the period, so that the interest payment amount is known in advance by both parties (the fixed rate is known of course, throughout the swap by both parties).

Although for the purposes of explaining swap structures, both parties are said to pay interest payments (and receive them), in practice only the net difference between both payments changes hands at the end of each interest payment. This eases the administration associated with swaps and reduces the number of cash flows for each swap. The party that is the net payer at the end of each period will make a payment to the counterparty. The first payment date will occur at the end of the first interest period, and subsequent payment dates will fall at the end of successive interest periods. The final payment date falls on the maturity date of the swap. The calculation of interest is given by (16.1):

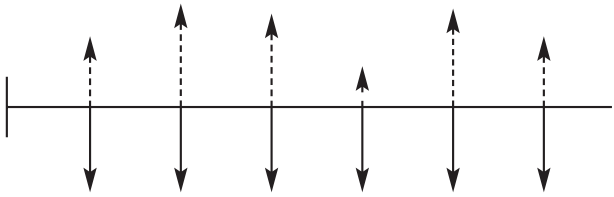
$$I = M \times r \times \frac{n}{B} \tag{16.1}$$

where I is the interest amount, M is the nominal amount of the swap and B is the interest day-base for the swap. Dollar and euro-denominated swaps use an actual/360 day-count, similar to other money market instruments in those currencies, while sterling swaps use an actual/365 day-count basis.

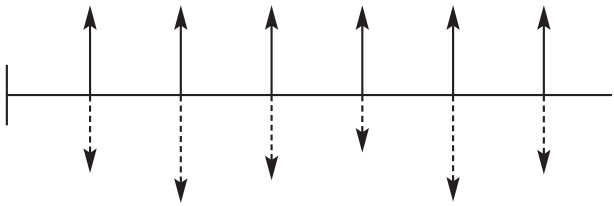
The cash flows resulting from a vanilla interest rate swap are illustrated in Figure 16.1, using the normal convention where cash inflows are shown as an arrow pointing up, while cash outflows are shown as an arrow pointing down. The counterparties in a swap transaction only pay across net cash flows, however, so at each interest payment date only one actual cash transfer will be made, by the net payer. This is shown as Figure 16.1(c).

Swap spreads and the swap yield curve

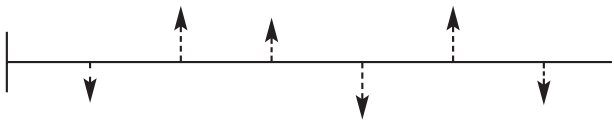
In the market, banks will quote two-way swap rates, on screens and on the telephone or via a dealing system such as Reuters. Brokers will also be active in relaying prices in the market. The convention in the market is for the swap market maker to set the floating leg at Libor and then quote the fixed rate that is payable



(a) Cash flows for fixed-rate payer



(b) Cash flows for floating-rate payer



(c) Net cash flows

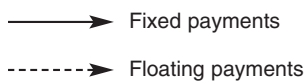


Figure 16.1 Cash flows for typical interest rate swap

for that maturity. So for a five-year swap a bank's swap desk might be willing to quote the following:

Floating-rate payer:	pay 6 mo. Libor
	receive fixed rate of 5.19%
Fixed-rate payer:	pay fixed rate of 5.25%
	receive 6 mo. Libor

In this case the bank is quoting an offer rate of 5.25%, which the fixed-rate payer will pay, in return for receiving Libor flat. The bid price quote is 5.19% which is

Table 16.2 Swap quotes

1 year	4.50	4.45	+17
2 year	4.69	4.62	+25
3 year	4.88	4.80	+23
4 year	5.15	5.05	+29
5 year	5.25	5.19	+31
10 year	5.50	5.40	+35

what a floating-rate payer will receive fixed. The bid–offer spread in this case is therefore 6 bps.

The fixed-rate quotes are always at a spread above the government bond yield curve. Let us assume that the five-year gilt is yielding 4.88%; in this case, then, the five-year swap bid rate is 31 bps above this yield. So the bank's swap trader could quote the swap rates as a spread above the benchmark bond yield curve, say 37–31, which is her *swap spread quote*. This means that the bank is happy to enter into a swap paying fixed 31 bps above the benchmark yield and receiving Libor, and receiving fixed 37 bps above the yield curve and paying Libor. The bank's screen on, say, Bloomberg or Reuters might look something like Table 16.2, which quotes the swap rates as well as the current spread over the government bond benchmark.

The swap spread is a function of the same factors that influence the spread over government bonds for other instruments. For shorter duration swaps, say up to three years, there are other yield curves that can be used in comparison, such as the cash market curve or a curve derived from futures prices. For longer-dated swaps the spread is determined mainly by the credit spreads that prevail in the corporate bond market. Because a swap is viewed as a package of long and short positions in fixed and floating-rate bonds, it is the credit spreads in these two markets that will determine the swap spread. This is logical; essentially it is the premium for greater credit risk involved in lending to corporates that dictates that a swap rate will be higher than the same maturity government bond yield. Technical factors will be responsible for day-to-day fluctuations in swap rates, such as the supply of corporate bonds and the level of demand for swaps, plus the cost to swap traders of hedging their swap positions.

In essence swap spreads over government bonds reflect the supply and demand conditions of both swaps and government bonds, as well as the market's view on the credit quality of swap counterparties. There is considerable information content in the swap yield curve, much like that in the government bond yield curve. During times of credit concerns in the market, such as the corrections in banking markets in the summer of 2007, the swap spread will increase, more so at longer maturities.

Zero-coupon swap pricing

Introduction

So far we have discussed how vanilla swap prices are often quoted as a spread over the benchmark government bond yield in that currency, and how this swap spread is mainly a function of the credit spread required by the market over the

government (risk-free) rate. This method is convenient and also logical because banks use government bonds as the main instrument when hedging their swap books. However because much bank swap trading is now conducted in non-standard, tailor-made swaps, this method can sometimes be unwieldy, as each swap needs to have its spread calculated to suit its particular characteristics. Therefore banks use a standard pricing method for all swaps known as *zero-coupon* swap pricing.

In Chapter 4 we referred to zero-coupon bonds and zero-coupon interest rates. Zero-coupon rates, or *spot rates*, are true interest rates for their particular term to maturity. In zero-coupon swap pricing, a bank will view all swaps, even the most complex, as a series of cash flows. The zero-coupon rates that apply now for each of the cash flows in a swap can be used to value these cash flows. Therefore to value and price a swap, each of the swap's cash flows are present-valued using known spot rates; the sum of these present values is the value of the swap.

In a swap the fixed-rate payments are known in advance and so it is straightforward to present-value them. The present value of the floating rate payments is usually estimated in two stages. First the implied forward rates can be calculated using (16.2). You should be quite familiar with this relationship from your reading of the earlier chapter.

$$rf_i = \left(\frac{df_i - 1}{df_{i+1}} \right) N \tag{16.2}$$

where

rf_i is the one-period forward rate starting at time i
df_i is the discount factor for the maturity period i
$df_i + 1$ is the discount factor for the period $i + 1$
N is the number of times per year that coupons are paid

By definition the floating-payment interest rates are not known in advance, so the swap bank will predict these, using the forward rates applicable to each payment date. The forward rates are those that are currently implied from spot rates. Once the size of the floating-rate payments has been estimated, these can also be valued by using the spot rates. The total value of the fixed and floating legs is the sum of all the present values, so the value of the total swap is the net of the present values of the fixed and floating legs.

While the term *zero-coupon* refers to an interest rate that applies to a discount instrument that pays no coupon and has one cash flow (at maturity), it is not necessary to have a functioning zero-coupon bond market in order to construct a zero-coupon yield curve. In practice most financial pricing models use a combination of the following instruments to construct zero-coupon yield curves:

- money market deposits
- interest-rate futures
- FRAs
- government bonds.

Frequently an overlap in the maturity period of all instruments is used; FRA rates are usually calculated from interest-rate futures so it is necessary to use only one of either FRA or futures rates.

Once a zero-coupon yield curve (*term structure*) is derived, this may be used to value a future cash flow maturing at any time along the term structure. This includes swaps: to price an interest-rate swap, we calculate the present value of each of the cash flows using the zero-coupon rates and then sum all the cash flows. As we noted above, while the fixed-rate payments are known in advance, the floating-rate payments must be estimated, using the forward rates implied by the zero-coupon yield curve. The net present value of the swap is the net difference between the present values of the fixed- and floating-rate legs.

Calculating the forward rate from spot rate discount factors

Remember that one way to view a swap is as a long position in a fixed-coupon bond that was funded at Libor, or against a short position in a floating-rate bond. The cash flows from such an arrangement would be paying floating-rate and receiving fixed-rate. In the former arrangement, where a long position in a fixed-rate bond is funded with a floating-rate loan, the cash flows from the principals will cancel out, as they are equal and opposite (assuming the price of the bond on purchase was par), leaving a collection of cash flows that mirror an interest-rate swap that pays floating and receives fixed. Therefore, the fixed rate on an interest-rate swap is the same as the coupon (and yield) on a bond priced at par, so that to calculate the fixed rate on an interest-rate swap is the same as calculating the coupon for a bond that we wish to issue at par.

The price of a bond paying semi-annual coupons is given by (16.3), which may be rearranged for the coupon rate r to provide an equation that enables us to determine the par yield, and hence the swap rate r, given by (16.4).

$$P = \frac{r_n}{2} df_1 + \frac{r_n}{2} df_2 + \cdots + \frac{r_n}{2} df_n + M df_n \tag{16.3}$$

where r_n is the coupon on an n-period bond with n coupons and M is the maturity payment. It can be shown then that:

$$\begin{aligned} r_n &= \frac{1 - df_n}{\frac{df_1}{2} + \frac{df_2}{2} + \cdots + \frac{df_n}{2}} \\ &= \frac{1 - df_n}{\sum_{i=1}^{n} \frac{df_i}{2}} \end{aligned} \tag{16.4}$$

For annual coupon bonds there is no denominator for the discount factor, while for bonds paying coupons on a frequency of N we replace the denominator 2 with N.[1]

[1] The expression also assumes an actual/365 day-count basis. If any other day-count convention is used, the 1/N factor must be replaced by a fraction made up of the actual number of days as the numerator and the appropriate year base as the denominator.

The expression at (16.4) may be rearranged again, using F for the coupon frequency, to obtain an equation which may be used to calculate the nth discount factor for an n-period swap rate, given at (16.5):

$$df_n = \frac{1 - r_n \sum_{i=1}^{n-1} \frac{df_i}{N}}{1 + \frac{rn}{N}} \tag{16.5}$$

The expression at (16.5) is the general expression for the *bootstrapping* process that we first encountered in Chapter 4. Essentially, to calculate the n-year discount factor we use the discount factors for the years 1 to n–1, and the n-year swap rate or zero-coupon rate. If we have the discount factor for any period, we may use (16.5) to determine the same period zero-coupon rate, after rearranging it, shown at (16.6):

$$rS_n = t_n \sqrt{\frac{1}{df_n} - 1} \tag{16.6}$$

Discount factors for spot rates may also be used to calculate forward rates. We know that:

$$df_1 = \frac{1}{\left(1 + \frac{rs_1}{N}\right)} \tag{16.7}$$

where rs is the zero-coupon rate. If we know the *forward rate* we may use this to calculate a second discount rate, shown by (16.8):

$$df_2 = \frac{df_1}{\left(1 + \frac{rf_1}{N}\right)} \tag{16.8}$$

where rf_1 is the forward rate. This is no use in itself; however we may derive from it an expression to enable us to calculate the discount factor at any point in time between the previous discount rate and the given forward rate for the period n to $n+1$, shown at (16.9), which may then be rearranged to give us the general expression to calculate a forward rate, given at (16.10).

$$df_{n+1} = \frac{df_n}{\left(1 + \frac{rf_n}{N}\right)} \tag{16.9}$$

$$rf_n = \left(\frac{df_n}{df_{n+1}} - 1\right) N \tag{16.10}$$

The general expression for an *n*-period discount rate at time *n* from the previous period forward rates is given by (16.11):

$$df_n = \frac{1}{\left(1+\frac{rf_{n-1}}{N}\right)} \times \frac{1}{\left(1+\frac{rf_{n-2}}{N}\right)} \times \cdots \times \frac{1}{\left(1+\frac{rf_0}{N}\right)} \tag{16.11}$$

$$df_n = \prod_{i=0}^{n-1} \left[\frac{1}{\left(1+\frac{rf_i}{N}\right)} \right]$$

From the above we may combine equations (16.4) and (16.10) to obtain the general expression for an *n*-period swap rate and zero-coupon rate, given by (16.12) and (16.13) respectively.

$$r_n = \frac{\sum_{i=1}^{n} \frac{rf_{i-1} df_i}{N}}{\sum_{i=1}^{n} \frac{df_i}{F}} \tag{16.12}$$

$$1 + rs_n = t_n \sqrt{\prod_{i=0}^{n-1} \left(1+\frac{rf_i}{n}\right)} \tag{16.13}$$

The two expressions tell us that the swap rate, which we have denoted as r_n, is shown by (16.12) to be the weighted average of the forward rates. A strip of FRAs would constitute an interest-rate swap, so a swap rate for a continuous period could be covered by a strip of FRAs. Therefore an average of the FRA rates would be the correct swap rate. As FRA rates are forward rates, we may be comfortable with (16.12), which states that the *n*-period swap rate is the average of the forward rates from rf_0 to rf_n.

To be accurate we must weight the forward rates, and these are weighted by the discount factors for each period. Note that although swap rates are derived from forward rates, interest payments under a swap are paid in the normal way at the end of an interest period, while payments for a FRA are made at the beginning of the period and must be discounted.

Equation (16.13) states that the zero-coupon rate is calculated from the geometric average of (one plus) the forward rates. Again, this is apparent from a reading of the case study example in Chapter 3. The *n*-period forward rate is obtained using the discount factors for periods *n* and *n*–1. The discount factor for the complete period is obtained by multiplying the individual discount factors together,

and exactly the same result would be obtained by using the zero-coupon interest-rate for the whole period to obtain the discount factor.[2]

Basic principles of an interest-rate swap

The rate charged on a newly transacted interest-rate swap is the one that gives its net present value as zero. The term *valuation* of a swap is used to denote the process of calculating the net present value of an existing swap, when marking-to-market the swap against current market interest rates. Therefore when we price a swap, we set its net present value to zero, while when we value a swap we set its fixed rate at the market rate and calculate the net present value.

To illustrate the basic principle, we price a plain vanilla interest rate swap with the terms set out below; for simplicity we assume that the annual fixed-rate payments are the same amount each year, although in practice there would be slight differences. Also assume we already have zero-coupon yields as shown in Table 16.3.

We use the zero-coupon rates to calculate the discount factors, and then use the discount factors to calculate the forward rates. This is done using equation (16.10). These forward rates are then used to predict what the floating-rate payments will be at each interest period. Both fixed-rate and floating-rate payments are then present-valued at the appropriate zero-coupon rate, which enables us to calculate the net present value.

The fixed-rate for the swap is calculated using equation (16.4) to give us:

$$\frac{1-0.71298618}{4.16187950}$$

or 6.8963%.

Nominal principal	£10,000,000
Fixed rate	6.8963%
Day count fixed	Actual/365
Day count floating	Actual/365
Payment frequency fixed	Annual
Payment frequency floating	Annual
Trade date	31 January 2000
Effective date	2 February 2000
Maturity date	2 February 2005
Term	Five years

[2] Zero-coupon and forward rates are also related in another way. If we change the zero-coupon rate rs_n and the forward rate rf_i into their continuously-compounded equivalent rates, given by $ln(1 + rs_n)$ and $ln(1 + rf_i)$, we may obtain an expression for the continuously-compounded zero-coupon rate as being the simple average of the continuously-compounded forward rates, given by:

$$rs'_n = \frac{1}{t_n}\sum_{i=0}^{n-1}\frac{rf'_i}{F}$$

Table 16.3 Generic interest-rate swap

Period	*Zero coupon rate %*	*Discount factor*	*Forward rate %*	*Fixed payment*	*Floating payment*	*PV fixed payment*	*PV floating payment*
1	5.5	0.947867298	5.5	689625	550000	653672.9858	521327.0142
2	6	0.88999644	6.502369605	689625	650236.9605	613763.7949	578708.58
3	6.25	0.833706493	6.751770257	689625	675177.0257	574944.8402	562899.4702
4	6.5	0.777323091	7.253534951	689625	725353.4951	536061.4366	563834.0208
5	7	0.712986179	9.023584719	689625	902358.4719	491693.094	643369.119
		4.161879501				2870137	2870137

Table 16.4 Generic interest-rate swap (Excel formulae)

Period	*Zero-coupon rate %*	*Discount factor*	*Forward rate %*	*Fixed payment*	*Floating payment*	*PV fixed payment*	*PV floating payment*
1	5.5	0.947867298	5.5	689625	"(F24*10000000)/100	"G24/1.055	"H24/(1.055)
2	6	0.88999644	"((E24/E25)-1)*100	689625	"(F25*10000000)/100	"G24/(1.06)^2	"H25/(1.06)^2
3	6.25	0.833706493	"((E25/E26)-1)*100	689625	"(F26*10000000)/100	"G24/(1.0625)^3	"H26/(1.0625^3)
4	6.5	0.777323091	"((E26/E27)-1)*100	689625	"(F27*10000000)/100	"G24/(1.065)^4	"H27/(1.065)^4
5	7	0.712986179	"((E27/E28)-1)*100	689625	"(F28*10000000)/100	"G24/(1.07)^5	"H28/(1.07)^5
		"SUM(E24:E28)				2,870,137	2,870,137

For reference the Microsoft Excel® formulae are shown in Table 16.4. It is not surprising that the net present value is zero, because the zero-coupon curve is used to derive the discount factors which are then used to derive the forward rates, which are used to value the swap. As with any financial instrument, the fair value is its breakeven price or hedge cost, and in this case the bank that is pricing the five-year swap shown in Table 16.3 could hedge the swap with a series of FRAs transacted at the forward rates shown. If the bank is paying fixed and receiving floating, the swap value to it will rise if there is a rise in market rates, and fall if there is a fall in market rates. Conversely, if the bank was receiving fixed and paying floating, the swap value to it would fall if there was a rise in rates, and vice versa.

This method is used to price any interest-rate swap, even exotic ones.

Valuation using final maturity discount factor

A short-cut to valuing the floating-leg payments of an interest-rate swap involves using the discount factor for the final maturity period. This is possible because, for the purposes of valuation, an exchange of principal at the beginning and end of the swap is conceptually the same as the floating-leg interest payments. This holds because, in an exchange of principal, the interest payments earned on investing the initial principal would be uncertain, as they are floating rate, while on maturity the original principal would be returned. The net result is a floating-rate level of receipts, exactly similar to the floating-leg payments in a swap. To value the principals then, we need only the final maturity discount rate.

To illustrate, consider Table 16.3, where the present value of both legs was found to be £2,870,137. The same result is obtained if we use the five-year discount factor, as shown below:

$$PV_{floating} - (10{,}000{,}000 \times 1) - (10{,}000{,}000 \times 0.71298618) = 2{,}870{,}137$$

The first term is the principal multiplied by the discount factor 1; this is because the present value of an amount valued immediately is unchanged (or rather, it is multiplied by the immediate payment discount factor, which is 1.0000).

Therefore we may use the principal amount of a swap if we wish to value the swap. This is of course for valuation only, as there is no actual exchange of principal in a swap.

Summary of IR swaps

A plain vanilla swap has the following characteristics:

- One leg of the swap is fixed-rate interest, while the other will be floating-rate, usually linked to a standard index such as Libor.
- The fixed rate is fixed through the entire life of the swap.
- The floating rate is set in advance of each period (quarterly, semi-annually or annually) and paid in arrears.
- Both legs have the same payment frequency.
- The maturity can be standard whole years up to 30 years, or set to match customer requirements.
- The notional principal remains constant during the life of the swap.

To meet customer demand banks can set up swaps that have variations on any or all of the above standard points. Some of the more common variations are discussed in the next section.

Non-vanilla interest-rate swaps

The swap market is very flexible, and instruments can be tailor-made to fit the requirements of individual customers. A wide variety of swap contracts have been traded in the market. Although the most common reference rate for the floating leg of a swap is six-month Libor, for a semi-annual paying floating leg other reference rates that have been used include three-month Libor, the prime rate (for dollar swaps), the one-month commercial paper rate, the Treasury bill rate and the municipal bond rate (again, for dollar swaps). The term of a swap need not be fixed; swaps may be *extendable* or *putable*. In an extendable swap, one of the parties has the right but not the obligation to extend the life of the swap beyond the fixed maturity date, while in a putable swap one party has the right to terminate the swap ahead of the specified maturity date.

It is also possible to transact options on swaps, known as *swaptions*. A swaption is the right to enter into a swap agreement at some point in the future, during the life of the option. Essentially a swaption is an option to exchange a fixed-rate bond cash flow for a floating-rate bond cash flow structure. As a floating-rate bond is valued on its principal value at the start of a swap, a swaption may be viewed as the value on a fixed-rate bond, with a strike price that is equal to the face value of the floating-rate bond.

A *constant maturity swap* is a swap in which the parties exchange a Libor rate for a fixed swap rate. For example, the terms of the swap might state that six-month Libor is exchanged for the five-year swap rate on a semi-annual basis for the next five years, or for the five-year government bond rate. In the US market the second type of constant maturity swap is known as a *constant maturity Treasury swap*.

Accreting and amortising swaps

In a plain vanilla swap the notional principal remains unchanged during the life of the swap. However it is possible to trade a swap where the notional principal varies during its life. An *accreting* (or *step-up*) swap is one in which the principal starts off at one level and then increases in amount over time. The opposite, an *amortising swap*, is one in which the notional reduces in size over time. An accreting swap would be useful where, for instance, a funding liability that is being hedged increases over time. The amortising swap might be employed by a borrower hedging a bond issue that featured sinking fund payments, where a part of the notional amount outstanding is paid off at set points during the life of the bond. If the principal fluctuates in amount, for example increasing in one year and then reducing in another, the swap is known as a *roller-coaster swap*.

Another application for an amortising swap is as a hedge for a loan that is itself an amortising one. Frequently this is combined with a forward-starting swap, to tie in with the cash flows payable on the loan. The pricing and valuation of an amortising swap is no different in principle from that for a vanilla interest-rate swap;

a single swap rate is calculated using the relevant discount factors, and at this rate the net present value of the swap cash flows will equal zero at the start of the swap.

Libor-in-arrears swap

In this type of swap (also known as a *back-set swap*) the setting date is just before the end of the accrual period for the floating-rate setting and not just before the start. Such a swap would be attractive to a counterparty who had a different view on interest rates compared to the market consensus. For instance in a rising yield curve environment, forward rates will be higher than current market rates, and this will be reflected in the pricing of a swap. A Libor-in-arrears swap would be priced higher than a conventional swap. If the floating-rate payer believed that interest rates would in fact rise more slowly than forward rates (and the market) were suggesting, she might wish to enter into an arrears swap as opposed to a conventional swap.

Basis swap

In a conventional swap one leg comprises fixed-rate payments and the other floating-rate payments. In a basis swap both legs are floating-rate, but linked to different money market indices. One leg is normally linked to Libor, while the other might be linked to the CD rate say, or the commercial paper rate. This type of swap would be used by a bank in the United States that had made loans that paid at the prime rate, and financed its loans at Libor. A basis swap would eliminate the *basis risk* between the bank's income and expense cash flows. Other basis swaps have been traded where both legs are linked to Libor, but at different maturities; for instance one leg might be at three-month Libor and the other at six-month Libor. In such a swap the basis is different and so is the payment frequency: one leg pays out semi-annually while the other would be paying on a quarterly basis. Note that where the payment frequencies differ, there is a higher level of counterparty risk for one of the parties. For instance, if one party is paying out on a monthly basis but receiving semi-annual cash flows, it would have made five interest payments before receiving one in return.

Margin swap

It is common to encounter swaps where there is a margin above or below Libor on the floating leg, as opposed to a floating leg of Libor flat. If a bank's borrowing is financed at Libor + 25 bps, it may wish to receive Libor + 25 bps in the swap so that its cash flows match exactly. The fixed rate quote for a swap must be adjusted correspondingly to allow for the margin on the floating side, so in our example if the fixed-rate quote is say, 6.00%, it would be adjusted to around 6.25%. Differences in the margin quoted on the fixed leg might arise if the day-count convention or payment frequency were to differ between fixed and floating legs. Another reason that there may be a margin is if the credit quality of the counterparty demanded it, so that highly rated counterparties may pay slightly below Libor, for instance.

Differential swap

A differential swap is a basis swap but with one of the legs calculated in a different currency. Typically one leg is floating-rate, while the other is floating-rate but with

the reference index rate for another currency, but denominated in the domestic currency. For example, a differential swap may have one party paying six-month sterling Libor, in sterling, on a notional principal of £10 million, and receiving euro-Libor, minus a margin, payable in sterling and on the same notional principal. Differential swaps are not very common and are the most difficult for a bank to hedge. The hedging is usually carried out using what is known as a *quanto* option.

Forward-start swap

A forward-start swap is one where the effective date is not the usual one or two days after the trade date but a considerable time afterwards, for instance say six months after the trade date. Such a swap might be entered into where one counterparty wanted to fix a hedge or cost of borrowing now, but for a point some time in the future. Typically this would be because the party considered that interest rates would rise or the cost of hedging would rise. The swap rate for a forward-starting swap is calculated in the same way as that for a vanilla swap.

Currency swaps

So far we have discussed swap contracts where the interest payments are both in the same currency. A *cross-currency swap*, or simply *currency swap*, is similar to an interest-rate swap, except that the currencies of the two legs are different. Like interest-rate swaps, the legs are usually fixed and floating-rate, although again it is common to encounter both fixed-rate or both floating-rate legs in a currency swap. On maturity of the swap there is an exchange of principals, and usually (but not always) there is an exchange of principals at the start of the swap. Where currencies are exchanged at the start of the swap, at the prevailing spot exchange rate for the two currencies, the exact amounts are exchanged back on maturity. During the time of the swap, the parties make interest payments in the currency that they have *received* where principals are exchanged. It may seem that exchanging the same amount on maturity gives rise to some sort of currency risk, but in fact it is this feature that removes any element of currency risk from the swap transaction.

Currency swaps are widely used in association with bond issues by borrowers who seek to tap opportunities in different markets but have no requirement for that market's currency. By means of a currency swap, a company can raise funds in virtually any market and swap the proceeds into the currency that it requires. Often the underwriting bank that is responsible for the bond issue will also arrange for the currency swap to be transacted. In a currency swap therefore, the exchange of principal means that the value of the principal amounts must be accounted for, and is dependent on the prevailing spot exchange rate between the two currencies.

Valuation of currency swaps

The same principles we established for the pricing and valuation of interest-rate swaps may be applied to currency swaps. A generic currency swap with fixed-rate payment legs would be valued at the fair value swap rate for each currency, which would give a net present value of zero. The cash flows are illustrated in Figure 16.2. This shows that the two swap rates in a fixed-fixed currency swap would be

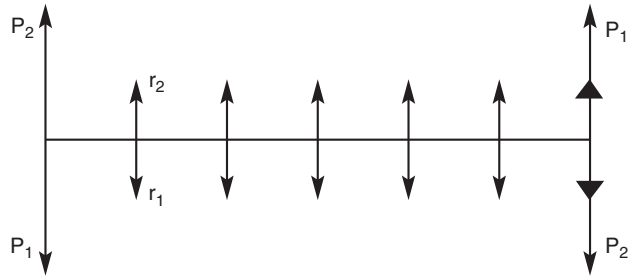


Figure 16.2 Fixed-fixed rate currency swap

identical to the same-maturity swap rate for each currency interest-rate swap. So the swap rates for a fixed-fixed five-year sterling/dollar currency swap would be the five-year sterling swap rate and the five-year dollar swap rate.

A floating-floating currency swap may be valued in the same way, and for valuation purposes the floating-leg payments are replaced with an exchange of principals, as we observed for the floating leg of an interest-rate swap. A fixed-floating currency swap is therefore valued at the fixed-rate swap rate for that currency for the fixed leg, and at Libor or the relevant reference rate for the floating leg.

Example 16.1: Bond issue and associated cross-currency swap

We illustrate currency swaps with an example from the market. A subsidiary of a US bank that invests in projects in the United States issues paper in markets around the world, in response to investor demand worldwide. The company's funding requirement is in US dollars; however it is active in issuing bonds in various currencies, according to where the most favourable conditions can be obtained. When an issue of debt is made in a currency other than dollars, the proceeds must be swapped into dollars for use in the United States, and interest payable on the swapped (dollar) proceeds. To facilitate this the issuer will enter into a currency swap. One of the bank's issues was a Swiss franc step-up bond, part of an overall Euro-MTN programme. The details of the bond are summarised below.

Issue date	March 1998
Maturity	March 2003
Size	CHF 15 million
Coupon	2.40% to 25 March 1999
	2.80% to 25 March 2000
	3.80% to 25 March 2001
	4.80% to 2 March 2002

The bond was also callable on each anniversary from March 1999 onwards, and in fact was called by the issuer at the earliest opportunity. The issuing

bank entered into a currency swap that resulted in the exchange of principals and the Swiss franc interest payments to be made by the swap counterparty; in return it paid US dollar three-month Libor during the life of the swap. At the prevailing spot rate on the effective date, CHF15 million was exchanged for US$10.304 million; these exact same amounts would be exchanged back on the maturity of the swap. When the issue was called the swap was cancelled and the swap counterparty paid a cancellation fee. The interest payment dates on the fixed leg of the swap matched the coupon dates of the bond exactly, as shown above. The floating leg of the swap paid USD Libor on a quarterly basis, as required by the bond issuer.

The structure is shown in Figure 16.3. A currency swap structure enables a bank or corporate to borrow money in virtually any currency in which a

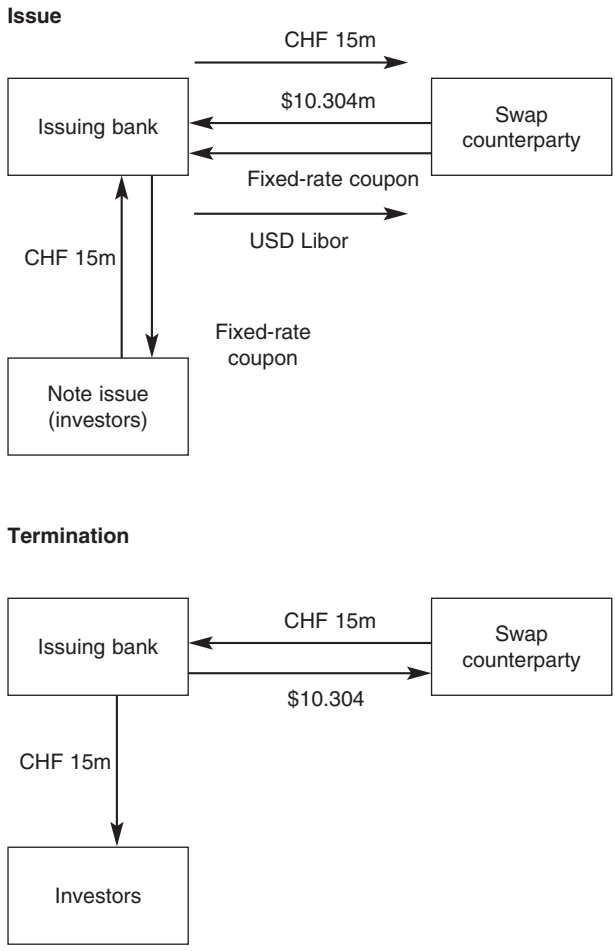


Figure 16.3 Bond issue with currency swap structure

liquid swap market exists, and swap this into a currency that is required. In our example the US bank was able to issue a bond that was attractive to investors. The swap mechanism also hedged the interest rate exposure on the Swiss franc note. The liability remaining for the issuer was quarterly floating rate interest on US dollars as part of the swap transaction.

Swaptions

Description

Swaptions are options on swaps. The buyer of a swaption has the right but not the obligation to enter into an interest-rate swap agreement during the life of the option. The terms of the swaption will specify whether the buyer is the fixed or floating-rate payer; the seller of the option (the *writer*) becomes the counterparty to the swap if the option is exercised. The convention is that if the buyer has the right to exercise the option as the fixed-rate payer, he has traded a *call swaption*, also known as a *payer swaption*, while if by exercising it the buyer of the swaption becomes the floating-rate payer, he has bought a *put swaption*, also known as a *receiver swaption*. The writer of the swaption is the party to the other leg.

Swaptions are similar to forward start swaps up to a point, but the buyer has the *option* of whether or not to commence payments on the effective date. A bank may purchase a call swaption if it expects interest rates to rise, and will exercise the option if indeed rates do rise as the bank has expected.

A company will use swaptions as part of an interest-rate hedge for a future exposure. For example, assume that a company will be entering into a five-year bank loan in three months' time. Interest on the loan is charged on a floating-rate basis, but the company intend to swap this to a fixed-rate liability after they have entered into the loan. As an added hedge, the company may choose to purchase a swaption that gives them the right to receive Libor and pay a fixed rate, say 10%, for a five-year period beginning in three months' time. When the time comes for the company to take out a swap and exchange their interest-rate liability in three months time (having entered into the loan), if the five-year swap rate is below 10%, the company will transact the swap in the normal way and the swaption will expire worthless. However if the five-year swap rate is above 10%, the company will instead exercise the swaption, giving it the right to enter into a five-year swap and paying a fixed rate of 10%.

Essentially the company has taken out protection to ensure that it does not have to pay a fixed rate of more than 10%. Hence swaptions can be used to guarantee a maximum swap rate liability. They are similar to forward-starting swaps, but do not commit a party to enter into a swap on fixed terms. The swaption enables a company to hedge against unfavourable movements in interest rates, but also to gain from favourable movements, although there is of course a cost associated with this, which is the premium paid for the swaption.

As with conventional put and call options, swaptions turn in-the-money under opposite circumstances. A call swaption increases in value as interest rates rise, and

a put swaption becomes more valuable as interest rates fall. Consider a one-year European call swaption on a five-year semi-annual interest-rate swap, purchased by a bank counterparty. The notional value is £10 million and the 'strike price' is 6%, against Libor. Assume that the price (premium) of the swaption is 25 bps, or £25,000. On expiry of the swaption, the buyer will either exercise it, in which case she will enter into a five-year swap paying 6% and receiving floating-rate interest, or elect to let the swaption expire with no value. If the five-year swap rate for counterparty of similar credit quality to the bank is above 6%, the swaption holder will exercise the swaption, while if the rate is below 6% the buyer will not exercise. The principle is the same for a put swaption, only in reverse.

Valuation

Swaptions are typically priced using the Black–Scholes or Black 76 option pricing models. These are used to value a European option on a swap, assuming that the appropriate swap rate at the expiry date of the option is lognormal. Consider a swaption with the following general terms:

Swap rate on expiry	*rs*
Swaption strike rate	*rX*
Maturity	*T*
Start date	*t*
Pay basis	*F* (say quarterly, semi-annual or annual)
Notional principal	*M*

If the actual swap rate on the maturity of the swaption is *rs*, the payoff from the swaption is given by:

$$\frac{M}{F}\max(r - r_n, 0)$$

The value of a swaption is essentially the difference between the strike rate and the swap rate at the time it is being valued. If a swaption is exercised, the payoff at each interest date is given by $(rs - rX) \times M \times F$. As a call swaption is only exercised when the swap rate is higher than the strike rate (that is, $rs > rX$), the option payoff on any interest payment in the swap is given by

$$Swaption_{InterestPayment} = \text{Max}[0,(rs - rX) \times \text{M} \times \text{F}] \tag{16.14}$$

It can then easily be shown that the value of a call swaption on expiry is given by

$$PV_{Swaption} = \sum_{n=1}^{N} Df_{(0,n)}(rs - rX) \times M \times F \tag{16.15}$$

where $Df_{(0,\,n)}$ is the spot rate discount factor for the term beginning now and ending at time *t*. By the same logic the value of a put swaption is given by the same expression except that $(rX - rs)$ is substituted at the relevant point above.

Pressing on, a swaption can be viewed as a collection of calls or puts on interest deposits or Libor, enabling us to use the Black model when valuing it. This means that we value each call or put on for a single payment in the swap, and then sum these payments to obtain the value of the swaption. The main assumption made when using this model is that the Libor rate follows a lognormal distribution over time, with constant volatility.

Consider a call swaption being valued at time t that matures at time T. We begin by valuing a single payment under the swap (assuming the option is exercised) made at time T_n. The point at time T_n is into the life of the swap, so that we have $T_n > T > t$. At the time of valuation, the option time to expiry is $T - t$ and there is $T_n - t$ until the nth payment. The value of this payment is given by

$$C_t = MFe^{-r(T_n - t)}[rsN(d_1) - rXN(d_2)] \tag{16.16}$$

where

C_t is the price of the call option on a single payment in the swap
r is the risk-free instantaneous interest rate
$N(.)$ is the cumulative normal distribution
σ is the interest-rate volatility

and where

$$d_1 = \frac{\text{In}(rs_t/rX) + \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}}$$

The remaining life of the swaption ($T - t$) governs the probability that it will expire in-the-money, determined using the lognormal distribution. On the other hand the interest payment itself is discounted (using $e^{-r(Tn-t)}$ over the period $T_n - t$ as it is not paid until time T_n).

Having valued a single interest payment, viewing the swap as a collection of interest payments, we value the call swaption as a collection of calls. Its value is given therefore by

$$PV_{Swaption} = \sum_{n=1}^{N} MFe^{-r(T_n - t)}[rsN(d_1) - rXN(d_2)] \tag{16.16}$$

where t, T and n are as before.

If we substitute discrete spot rate discount factors instead of the continuous form given by (16.16) the expression becomes

$$PV_{Swaption} = MF[rsN(d_1) - rXN(d_2)]\sum_{n=1}^{N} Df_{t,Tn} \tag{16.17}$$

Example 16.2: Swaption pricing[3]

We present a new term structure environment in this example to illustrate the basic concepts. This is shown in Table 16.5. We wish to price a forward-starting annual interest swap starting in two years for a term of three years. The swap has a notional of £10 million.

Table 16.5 Interest rate data for swaption valuation

Date	*Term (years)*	*Discount factor*	*Par yield*	*Zero-coupon rate*	*Forward rate*
18 Feb 2001	0	1	5	5	5
18 Feb 2002	1	0.95238095	5.00	5	6.03015
18 Feb 2003	2	0.89821711	5.50	5.51382	7.10333
18 Feb 2004	3	0.83864539	6.00	6.04102	6.66173
18 Feb 2005	4	0.78613195	6.15	6.19602	6.71967
20 Feb 2006	5	0.73637858	6.25	6.30071	8.05230
19 Feb 2007	6	0.68165163	6.50	6.58946	8.70869
18 Feb 2008	7	0.62719194	6.75	6.88862	9.40246
18 Feb 2009	8	0.57315372	7.00	7.20016	10.18050
18 Feb 2010	9	0.52019523	7.25	7.52709	5.80396
18 Feb 2011	10	0.49165950	7.15	7.35361	6.16366

The swap rate is given by:

$$rs = \frac{\sum_{t=1}^{N} rf_{(t-1)t} \cdot Df_{0,t}}{\sum_{t=1}^{N} Df_{0,t}}$$

where rf is the forward rate.

Using the above expression, the numerator in this example is

$$(0.0666 \times 0.8386) + (0.0672 \times 0.7861) + (0.0805 \times 0.7364) \text{ or } 0.1701$$

The denominator is

$$0.8386 + 0.7861 + 0.7634 \text{ or } 2.3881$$

Therefore the forward-starting swap rate is 0.1701/2.3881 or 0.071228 (7.123%).

We now turn to the call swaption on this swap, the buyer of which acquires the right to enter into a three-year swap paying the fixed-rate swap

[3] This example follows the approach described in Kolb (2000), although here we use discount factors in the calculation whereas Kolb uses *zero-coupon factors* which are (1 + spot rate).

rate of 7.00%. If the volatility of the forward swap rate is 0.20, the d_1 and d_2 terms are

$$d_1 = \frac{\text{In}\left(\frac{rs}{rX}\right) + \frac{\sigma^2}{2}(T-t)}{\sigma\sqrt{T-t}} = \frac{\text{In}\left(\frac{0.071228}{0.07}\right) + \left(\frac{0.2^2}{2} \times 2\right)}{0.2\sqrt{2}}$$

or 0.2029068

$$d_2 = d_1 - \sigma\sqrt{T-t} = 0.20290618 - 0.2(1.4142) \text{ or } -0.079934.$$

The cumulative normal values are

$N(d_1)$ + $N(0.2029)$ which is 0.580397
$N(d_2)$ + N(−0.079934) which is 0.468145.[4]

From above we know that $\Sigma Df_{t,T_n}$ is 2.3881. So using (16.17) we calculate the value of the call swaption to be:

$$PV_{Swaption_t} = MF[rsN(d_1) - rXN(d_2)]\sum_{n=1}^{N} Df_{t,Tn}$$

= 10,000,000 × 1 × [0.07228 × 0.580397 − 0.07 × 0.468145] × 2.3881
= 219,250

Option premiums are frequently quoted as basis points of the notional amount, so in this case the premium is (219,250 / 10,000,000) or 219.25 bps.

An overview of interest-rate swap applications

In this section we review some of the principal uses of swaps as a hedging tool.

Corporate applications

Swaps are part of the OTC market, and so they can be tailored to suit the particular requirements of the user. It is common for swaps to be structured so that they match particular payment dates, payment frequencies and Libor margins, which may characterise the underlying exposure of the customer. As the market in interest-rate swaps is so large, liquid and competitive, banks are willing to quote rates and structure swaps for virtually all customers.

Swap applications can be viewed as being of two main types, asset-linked swaps and liability-linked swaps. *Asset-linked swaps* are created when the swap is linked to an asset such as a bond in order to change the characteristics of the income

[4] These values may be found from standard normal distribution tables or using the Microsoft Excel formula = NORMSDIST().

stream for investors. *Liability-linked swaps* are traded when borrowers of funds wish to change the pattern of their cash flows. Of course, just as with repo transactions, the designation of a swap in such terms depends on from whose point of view one is looking at the swap. An *asset-linked swap hedge* is a liability-linked hedge for the counterparty, except in the case of swap market-making banks which make two-way quotes in the instruments.

A straightforward application of an interest-rate swap is when a borrower wishes to convert a floating-rate liability into a fixed-rate one, usually in order to remove the exposure to upward moves in interest rates. For instance a company may wish to fix its financing costs. Let us assume a company currently borrowing money at a floating rate, say six month Libor +100 bps, fears that interest rates may rise in the remaining three years of its loan. It enters into a three-year semi-annual interest rate swap with a bank, as the fixed-rate payer, paying say 6.75% against receiving six-month Libor. This fixes the company's borrowing costs for three years at 7.75% (7.99% effective annual rate). This is shown in Figure 16.4.

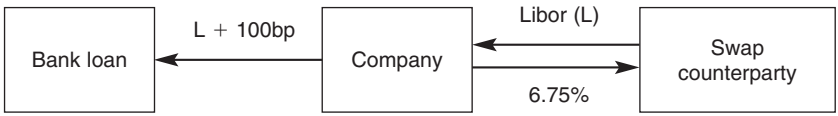


Figure 16.4 Changing liability from floating to fixed-rate

Example 16.3: Liability-linked swap, fixed to floating to fixed-rate exposure

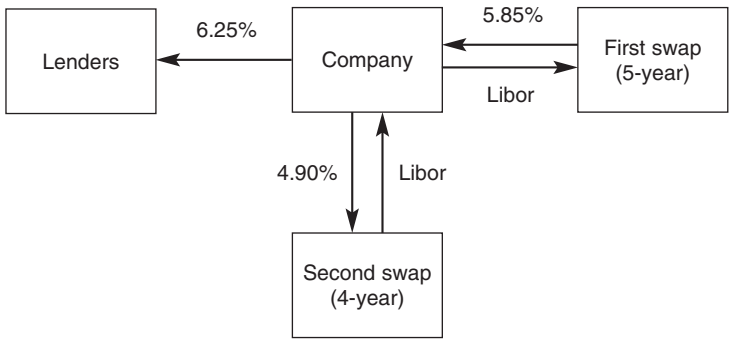


Figure 16.5 Liability-linked swap, fixed to floating to fixed-rate exposure

A corporate borrows for five years at a rate of 6.25% and shortly after enters into a swap paying floating rate, so that its net borrowing cost is Libor + 40 bps. After one year swap rates have fallen such that the company is quoted four-year swap rates as 4.90–84%. The company decides to switch

back into fixed-rate liability in order to take advantage of the lower interest rate environment. It enters into a second swap paying fixed at 4.90% and receiving Libor. The net borrowing cost is now 5.30%. The arrangement is illustrated in Figure 16.5. The company has saved 95 bps on its original borrowing cost, which is the difference between the two swap rates.

Asset-linked swap structures might be required when for example, investors require a fixed-interest security when floating-rate assets are available. Borrowers often issue FRNs, the holders of which might prefer to switch the income stream into fixed coupons. As an example, consider a local authority pension fund holding two-year floating-rate gilts. This is an asset of the highest quality, paying Libid minus 12.5 bps. The pension fund wishes to swap the cash flows to create a fixed-interest asset. It obtains a quote for a tailor-made swap where the floating leg pays Libid, the quote being 5.55–50%. By entering into this swap the pension fund has in place a structure that pays a fixed coupon of 5.375%. This is shown in Figure 16.6.

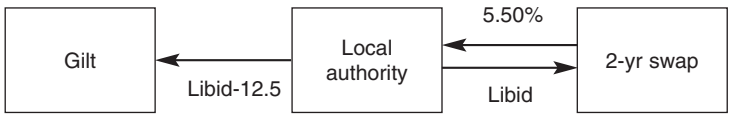


Figure 16.6 Transforming a floating-rate asset to fixed-rate

Hedging bond instruments using interest-rate swaps

We illustrate here a generic approach to the hedging of bond positions using interest-rate swaps. The bond trader has the option of using other bonds, bond futures or bond options, as well as swaps, when hedging the interest-rate risk exposure of a bond position. However swaps are particularly efficient instruments to use because they display positive convexity characteristics, that is, the increase in value of a swap for a fall in interest rates exceeds the loss in value with a similar magnitude rise in rates. This is exactly the price/yield profile of vanilla bonds.

The primary risk measure we require when hedging using a swap is its present value of a basis point or PVBP.[5] This measures the price sensitivity of the swap for a basis point change in interest rates. The PVBP measure is used to calculate the hedge ratio when hedging a bond position. The PVBP can be given by:

$$PVBP = \frac{\text{Change in swap value}}{\text{Rate change in basis points}} \tag{16.18}$$

which can be written as

$$PVBP = \frac{dS}{dr} \tag{16.19}$$

[5] This is also known as DVBP or dollar value of a basis point in the US market.

Using the basic relationship for the value of a swap, which is viewed as the difference between the values of a fixed coupon bond and equivalent-maturity floating-rate bond (see Table 16.1) we can also write:

$$PVBP = \frac{d\text{Fixed bond}}{dr} - \frac{d\text{Floating bond}}{dr} \tag{16.20}$$

which essentially states that the basis point value of the swap is the difference in the basis point values of the fixed-coupon and floating-rate bonds. The value is usually calculated for a notional £1 million of swap. The calculation is based on the duration and modified duration calculations used for bonds (see Chapter 2), and assumes that there is a parallel shift in the yield curve.

Table 16.6 illustrates how equations (16.19) and (16.20) can be used to obtain the PVBP of a swap. Hypothetical five-year bonds are used in the example. The PVBP for a bond can be calculated using Bloomberg or the MDURATION function on Microsoft Excel. Using either of the two equations above we see that the PVBP of the swap is £425.00. This is shown below.

Calculating the PVBP using (16.19) we have:

$$PVBP_{swap} = \frac{dS}{dr} = \frac{4624 - (-4236)}{20} = 425$$

while using (16.20) we obtain the same result using the bond values:

$$PVBP_{swap} = PVBP_{fixed} - PVBP_{floating}$$

$$= \frac{1{,}004{,}940 - 995{,}171}{20} - \frac{1{,}000{,}640 - 999{,}371}{20} = 488.45 - 63.45 = 425.00$$

The swap basis point value is lower than that of the five-year fixed-coupon bond, that is, £425 compared to £488.45. This is because of the impact of the floating-rate bond risk measure, which reduces the risk exposure of the swap as a whole by £63.45. As a rough rule of thumb, the PVBP of a swap is approximately equal to

Table 16.6 PVBP for interest-rate swap

Term to maturity	5 years
Fixed leg	6.50%
Basis	Semi-annual, act/365
Floating leg	6-month Libor
Basis	Semi-annual, act/365
Nominal amount	£1,000,000

	Present value £ Rate change − 10 bps	*0 bps*	*Rate change + 10 bps*
Fixed coupon bond	1,004,940	1,000,000	995,171
Floating rate bond	1,000,640	1,000,000	999,371
Swap	4,264	0	4,236

that of a fixed-rate bond that has a maturity similar to the period from the next coupon reset date of the swap through to the maturity date of the swap. This means that a 10-year semi-annual paying swap would have a PVBP close to that of a 9.5-year fixed-rate bond, and a 5.50-year swap would have a PVBP similar to that of a 5-year bond.

When using swaps as hedge tools, we bear in mind that over time the PVBP of swaps behaves differently from that of bonds. Immediately preceding an interest reset date the PVBP of a swap will be near-identical to that of the same-maturity fixed-rate bond, because the PVBP of a floating-rate bond at this time has essentially nil value. Immediately after the reset date the swap PVBP will be near-identical to that of a bond that matures at the next reset date. This means that at the point (and this point only) right after the reset, the swap PVBP will decrease by the amount of the floating-rate PVBP. In between reset dates the swap PVBP is quite stable, as the effects of the fixed and floating-rate PVBP changes cancel each other out. Contrast this with the fixed-rate PVBP, which decreases in value over time in a stable fashion.[6] This feature is illustrated in Figure 16.7. A slight anomaly is that the PVBP of a swap actually increases by a small amount between reset dates; this is because the PVBP of a floating-rate bond decreases at a slightly faster rate than that of the fixed-rate bond during this time.

Hedging bond instruments with interest-rate swaps is conceptually similar to hedging with another bond or with bond futures contracts. If one is holding a long

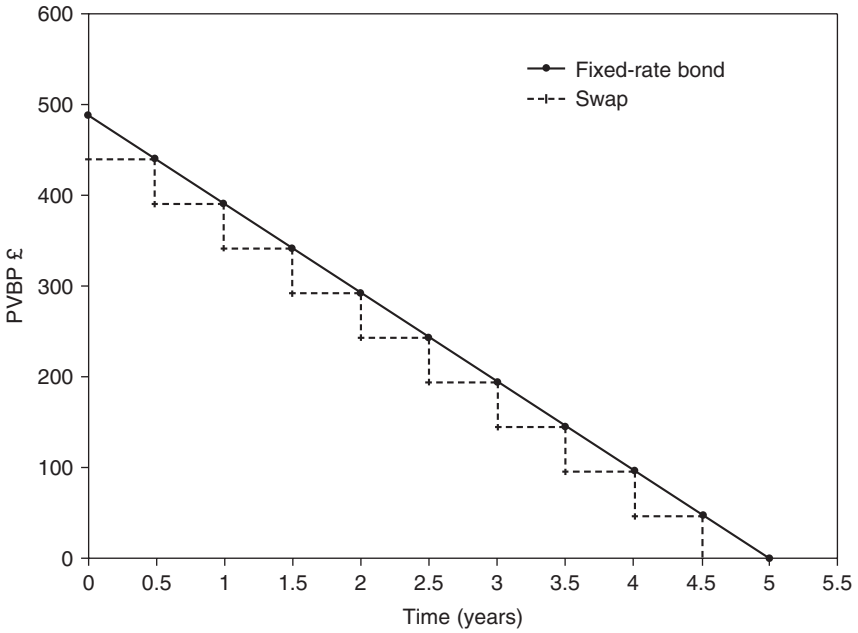


Figure 16.7 PVBP of a 5-year swap and fixed-rate bond maturity period

[6] This assumes no large-scale sudden yield movements.

position in a vanilla bond, the hedge requires a long position in the swap: remember that a long position in a swap is to be paying fixed (and receiving floating). This hedges the receipt of fixed from the bond position. The change in the value of the swap will match the change in value of the bond, only in the opposite direction.[7] The maturity of the swap should match as closely as possible that of the bond. As swaps are OTC contracts, it should be possible to match interest dates as well as maturity dates. If one is short the bond, the hedge is to be short the swap, so the receipt of fixed matches the pay-fixed liability of the short bond position.

The correct nominal amount of the swap to put on is established using the PVBP hedge ratio. This is given as:

$$\text{Hedge ratio} = \frac{PVBP_{bond}}{PVBP_{swap}} \tag{16.21}$$

This technique is still used in the market, but suffers from the assumption of parallel yield curve shifts, and can therefore lead to significant hedging error at times. More advanced techniques are used by banks when hedging books using swaps, but space does not permit any discussion of them here. Some of these techniques are discussed in the author's forthcoming book on swaps.

Selected bibliography and references

Bicksler, J. and Chen, A. 'An economic analysis of interest rate swaps', *Journal of Finance*, 41, 3, 1986, pp. 645–55.

Brotherton-Ratcliffe, R. and Iben, B. 'Yield curve applications of swap products', in Schwartz, R. and Smith, C. (eds), *Advanced Strategies in Financial Risk Management*, New York Institute of Finance, 1993.

Choudhry, M. *Bond Market Securities*, FT Prentice Hall 2001, ch. 11.

Das, S. *Swaps and Financial Derivatives*, 2nd edn, IFR Publishing, 1994.

Decovny, S. *Swaps*, 2nd edn, FT Prentice Hall, 1998.

Dunbar, N. 'Swaps volumes see euro wane', *Risk*, September 2000.

Eales, B. *Financial Risk Management*, McGraw Hill, 1995, ch. 3.

Fabozzi, F. (ed.), *Perspectives on Interest Rate Risk Management for Money Managers and Traders*, FJF Associates, 1998.

Gup, B. and Brooks, R. *Interest Rate Risk Management*, Irwin, 1993.

Henna, P. *Interest-Rate Risk Management using Futures and Swaps*, Probus, 1991.

International Swaps and Derivatives Association, *Code of Standard Working, Assumptions and Provisions for Swaps*, New York, 1991.

Jarrow, R. and Turnbull, S. *Derivative Securities*, 2nd edn, South-Western, 2000.

Khan, M. 'Online platforms battle for business', *Risk*, September 2000.

Kolb, R. *Futures, Options and Swaps,* 3rd edn, Blackwell, 2000.

Li, A. and Raghavan, V. R. 'LIBOR-in-arrears swaps', *Journal of Derivatives*, 3, Spring 1996, pp. 44–8.

Lindsay, R. 'High wire act', *Risk,* August 2000.

Marshall, J. and Kapner, K. *Understanding Swap Finance,* South-Western Publishing, 1990.

Turnbull, S. 'Swaps: a zero sum game', *Financial Management*, 16, Spring 1987, pp. 15–21.

[7] The change will not be an exact mirror. It is very difficult to establish a precise hedge for a number of reasons, which include differences in day-count bases, maturity mismatches and basis risk.

17
Options I

As a risk management tool, options allow banks and corporates to hedge market exposure but also to gain from upside moves in the market; this makes them unique amongst hedging instruments. Options have special characteristics that make them stand apart from other classes of derivatives. As they confer a right to conduct a certain transaction, but not an obligation, their payoff profile is different from other financial assets, both cash and off-balance sheet (OBS). This makes an option more of an insurance policy rather than a pure hedging instrument, as the person who has purchased the option for hedging purposes need only exercise it if required. The price of the option is in effect the insurance premium that has been paid for peace of mind.

Of course options are also used for purposes other than hedging. They are used as part of speculative and arbitrage trading, and option market makers generate returns from profitably managing the risk on their option books.

The range of combinations of options that can be dealt today, and the complex structured products that they form part of, are constrained only by imagination and customer requirements. Virtually all participants in capital markets will have some requirement that can be met by the use of options. The subject is a large one, and there are a number of specialist texts devoted to them. In this chapter we introduce the basics of options; subsequent chapters review option pricing, the main sensitivity measures used in running an option book, and the uses to which options may be put. Key reference articles and publications are also listed in the bibliography.

Introduction

An *option* is a contract in which the buyer has the right, but not the obligation, to buy or sell an underlying asset at a predetermined price during a specified period of time. The seller of the option, known as the *writer*, grants this right to the buyer in return for receiving the price of the option, known as the *premium*. An option that grants the right to buy an asset is a *call option*, while the corresponding right to sell an asset is a *put option*. The option buyer has a long position in the option and the option seller has a short position in the option.

Before looking at the other terms that define an option contract, we discuss the main feature that differentiates an option from all other derivative instruments, and from cash assets. Because options confer on a buyer the right to effect a transaction, but not the obligation (and correspondingly on a seller the obligation, if requested by the buyer, to effect a transaction), their risk/reward characteristics are different from other financial products. The payoff profile from holding an option is unlike that of any other instrument.

Let us consider the payoff profiles for a vanilla call option and a gilt futures contract. Suppose that a trader buys one lot of the gilt futures contract at 114.00 and holds it for one month before selling it. On closing the position, the profit made will depend on the contract sale price. If it is above 114.00 the trader will have made a profit and if below 114.00 she will have made a loss. On one lot this represents a £1000 gain for each point above 114.00. The same applies to someone who had a short position in the contract and closed it out – if the contract is bought back at any price below 114.00 the trader will realise a profit. The profile is shown in Figure 17.1.

This profile is the same for other derivative instruments such as FRAs and swaps, and of course for cash instruments such as bonds or equity. The payoff profile therefore has a *linear* characteristic, and it is linear whether one has bought or sold the contract.

The profile for an option contract differs from the conventional one. Because options confer a right to one party but not an obligation (the buyer), and an obligation but not a right to the seller, the profile will differ according to whether one is the buyer or seller. Suppose now that our trader buys a call option that grants the right to buy a gilt futures contract at a price of 114.00 at some point

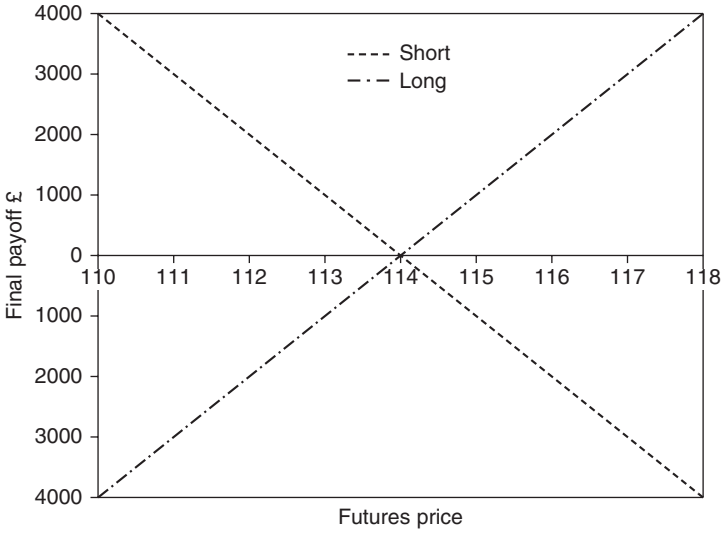


Figure 17.1 Payoff profile for a bond futures contract

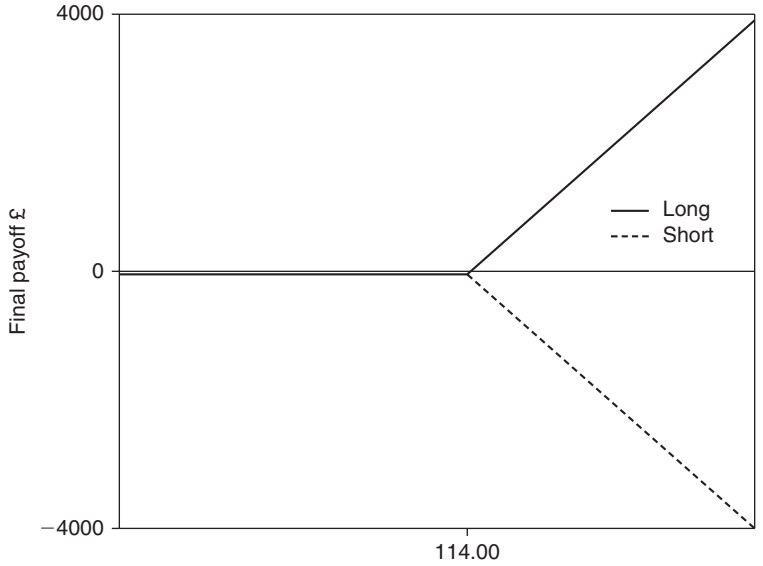


Figure 17.2 Payoff profile for call option contract

during the life of the option. Her resulting payoff profile will be like that shown in Figure 17.2.

If during the life of the option, the price of the futures contract rises above 114.00, the trader will exercise her right to buy the future, under the terms of the option contract. This is known as *exercising* the option. If on the other hand the price of the future falls below 114.00, the trader will not exercise the option, and unless there is a reversal in price of the future, it will eventually expire worthless, on its maturity date. In this respect it is exactly like an equity or bond warrant.

The seller of this particular option has a very different payout profile. If the price of the future rises above 114.00 and the option is exercised, the seller will bear the loss equal to the profit that the buyer is now benefiting from. The seller's payoff profile is also shown in Figure 17.2, as the dashed line. If the option is not exercised and expires, for the seller the trade will have generated premium income, which is revenue income that contributes to the profit and loss account.

This illustrates how unlike every other financial instrument, the holders of long and short positions in options do not have the same symmetrical payoff profile. The buyer of the call option will benefit if the price of the underlying asset rises, but will not lose if the price falls (except the funds paid for purchasing the rights under the option). The seller of the call option will suffer loss if the price of the underlying asset rises, but will not benefit if it falls (except realising the funds received for writing the option). The buyer has a right but not an obligation, while the seller has an obligation if the option is exercised. The premium charged for the option is the seller's compensation for granting such a right to the buyer.

Let us recap on the basic features of the call option. A call option is the right to buy, without any obligation, a specified quantity of the underlying asset at a given price on or before the expiry date of the option. A long position in a call option allows the holder, as shown in Figure 17.2, to benefit from a rise in the market price of the underlying asset. If our trader wanted to benefit from a fall in the market level, but did not want to short the market, she would buy a *put* option. A put option is the right to sell, again without any obligation, a specified quantity of the underlying asset at a given price on or before the expiry date of the option. Put options have the same payoff profile as call options, but in the opposite direction. Remember also that the payoff profile is different for the buyer and seller of an option. The buyer of a call option will profit if the market price of the underlying asset rises, but will not lose if the price falls (at least, not with regard to the option position). The writer of the option will not profit whatever direction the market moves in, and will lose if the market rises. The compensation for taking on this risk is the premium paid for writing the option, which is why we likened options to insurance policies at the start of the chapter.

Originally options were written on commodities such as wheat and sugar. Nowadays these are referred to as *options on physicals*, while options on financial assets are known as *financial options*. Today one is able to buy or sell an option on a wide range of underlying instruments, including financial products such as foreign exchange, bonds, equities and commodities, and derivatives such as futures, swaps, equity indices and other options.

Option terminology

Let us now consider the basic terminology used in the options markets.

A *call* option grants the buyer the right to buy the underlying asset, while a *put* option grants the buyer the right to sell the underlying asset. There are therefore four possible positions that an option trader may put on, long a call or put, and short a call or put. The payoff profiles for each type are shown in Figure 17.3.

The *strike price* describes the price at which an option is exercised. For example a call option to buy ordinary shares of a listed company might have a strike price of £10.00. This means that if the option is exercised, the buyer will pay £10 per share.

Options are generally either *American* or *European* style, which defines the times during the option's life when it can be exercised. There is no geographic relevance to these terms, as both styles can be traded in any market. There is also another type, *Bermudan* style options, which can be exercised at pre-set dates.[1] For reasons that we shall discuss later, it is very rare for an American option to be exercised ahead of its expiry date, so this distinction has little impact in practice, although of course the pricing model being used to value European options must be modified to handle American options. The holder of a European option cannot exercise it prior to expiry; however if she wishes to realise its value she will sell it in the market.

[1]We are told because Bermuda is midway between Europe and America. A colleague also informs us that it is 'Asian' for average-rate options because these originated in Japanese commodity markets.

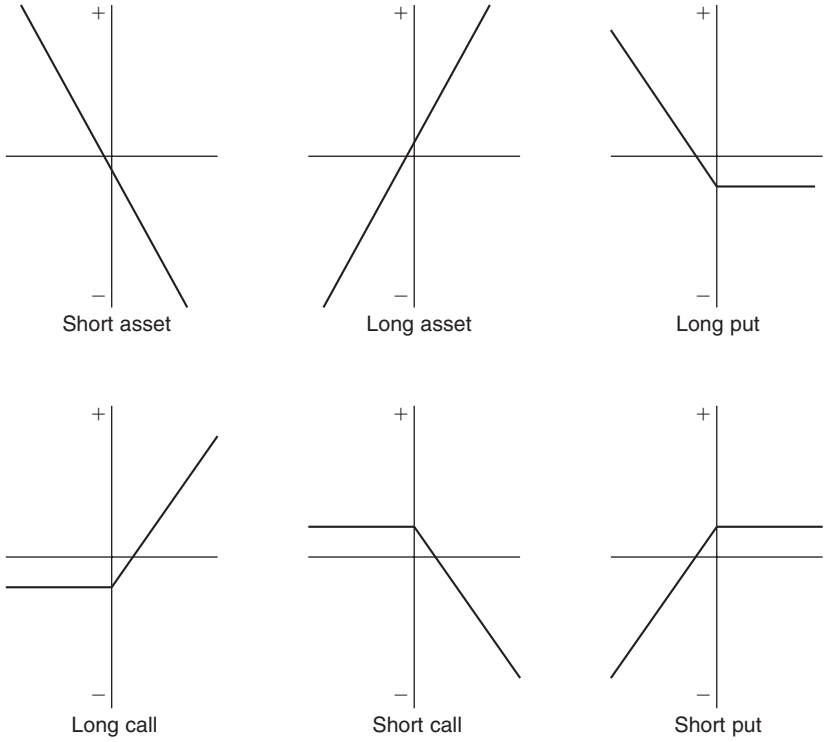


Figure 17.3 Basic option payoff profiles

The *premium* of an option is the price at which the option is sold. Option premium is made up of two constituents, *intrinsic value* and *time value*.

The intrinsic value of an option is the value of the option if it is exercised immediately, and it represents the difference between the strike price and the current underlying asset price. If a call option on a bond futures contract has a strike price of 100.00 and the future is currently trading at 105.00, the intrinsic value of the option is 5.00, as this would be the immediate profit gain to the option holder if it were exercised. Since an option will only be exercised if there is benefit to the holder from so doing, its intrinsic value will never be less than zero. So in our example if the bond future was trading at 95.00 the intrinsic value of the call option would be zero, not –5.00. For a put option the intrinsic value is the amount by which the current underlying price is below the strike price. When an option has intrinsic value it is described as being *in-the-money*. When the strike price for a call option is higher than the underlying price (or for a put option is lower than the underlying price) and has no intrinsic value it is said to be *out-of-the-money*. An option for which the strike price is equal to the current underlying price is said to be *at-the-money*. This term is normally used at the time the option is first traded, in cases where the strike price is set to the current price of the underlying asset.

Table 17.1 Basic option terminology

Call	The right to buy the underlying asset
Put	The right to sell the underlying asset
Buyer	The person who has purchased the option and has the right to exercise it if she wishes
Writer	The person who has sold the option and has the obligation to perform if the option is exercised
Strike price	The price at which the option may be exercised, also known as the exercise price
Expiry date	The last date on which the option can be exercised, also known as the maturity date
American	The style of option; an American option can be exercised at any time up to the expiry date
European	An option which may be exercised on the maturity date only, and not before
Premium	The price of the option, paid by the buyer to the seller
Intrinsic value	The value of the option if was exercised today, which is the difference between the strike price and the underlying asset price
Time value	The difference between the current price of the option and its intrinsic value
In-the-money	The term for an option that has intrinsic value
At-the-money	An option for which the strike price is identical to the underlying asset price
Out-of-the-money	An option that has no intrinsic value

The *time value* of an option is the amount by which the option value exceeds the intrinsic value. An option writer will almost always demand a premium that is higher than the option's intrinsic value, because of the risk that the writer is taking on. This reflects the fact that over time the price of the underlying asset may change sufficiently to produce a much higher intrinsic value. During the life of an option, the option writer has nothing more to gain over the initial premium at which the option was sold; however until expiry there is a chance that the writer will lose if the markets move against her, hence the inclusion of a time value element. The value of an option that is out-of-the-money is composed entirely of time value.

Table 17.1 summarises the main option terminology that we have just been discussing.

Option instruments

Options are traded both on recognised exchanges and in the OTC market. The primary difference between the two types is that exchange-traded options are standardised contracts and essentially plain vanilla instruments, while OTC options can take on virtually any shape or form. Options traded on an exchange

are often options on a futures contract, so for example a gilt option on LIFFE in London is written on the exchange's gilt futures contract. The exercise of a futures option will result in a long position in a futures contract being assigned to the party that is long the option, and a short position in the future to the party that is short the option. Note that exchange-traded options on US Treasuries are quoted in option ticks that are half the bond tick, that is 1/64th rather than 1/32nd. The same applied to gilt options on LIFFE until gilts themselves switched to decimal pricing at the end of 1998.

Like OTC options, those traded on an exchange can be either American or European style. For example on the Philadelphia Currency Options Exchange both versions are available, although on LIFFE most options are American style. Exchange-traded options are available on the following:

- *Ordinary shares*. Major exchanges including the New York Stock Exchange, LIFFE, Eurex, the Chicago Board Options Exchange (CBOE), and SIMEX in Singapore trade options on corporation ordinary shares.
- *Options on futures*. Most exchanges trade an option contract written on the futures that are traded on the exchange, which expires one or two days before the futures contract itself expires. In certain cases such as those traded on the Philadelphia exchange, cash settlement is available, so that if for example the holder of a call exercises, she will be assigned a long position in the future as well as the cash value of the difference between the strike price and the futures price. One of the most heavily traded exchange-traded options contracts is the Treasury bond option, written on the futures contract traded on the Chicago Board of Trade options exchange.
- *Stock index options*. These are equity market instruments that are popular for speculating and hedging, for example the FTSE-100 option on LIFFE and the S&P500 on CBOE. Settlement is in cash and not the shares that constitute the underlying index, much like the settlement of an index futures contract.
- *Bond options*. Options on bonds are invariably written on the bond futures contract, for example the aforementioned Treasury bond option or LIFFE's gilt option. Options written on the cash bond must be traded in the OTC market.
- *Interest-rate options*. These are also options on futures, as they are written on the exchange's 90-day interest-rate futures contract;
- *Foreign currency options*. This is rarer among exchange-traded options, and the major exchange is in Philadelphia. Its sterling option contract for example is for an underlying amount of £31,250.

Option trading on an exchange is similar to that for futures, and involves transfer of margin on a daily basis. Individual exchanges have their own procedures; for example on LIFFE the option premium is effectively paid via the variation margin. The amount of variation margin paid or received on a daily basis for each position reflects the change in the price of the option. So if for example an option were to expire on maturity with no intrinsic value, the variation margin payments made

during its life would be equal to the change in value from the day it was traded to zero. The option trader does not pay a separate premium on the day of the day the position is put on. On certain other exchanges, though, it is the other way around, and the option buyer will pay a premium on the day of purchase but then pay no variation margin. Some exchanges allow traders to select either method. Margin is compulsory for a party that writes options on the exchange.

The other option market is the OTC market, where there is a great variety of different instruments traded. As with products such as swaps, the significant advantage of OTC options is that they can be tailored to meet the specific requirements of the buyer. Hence they are ideally suited as risk management instruments for corporate and financial institutions, because they can be used to structure hedges that match perfectly the risk exposure of the buying party. Some of the more ingenious structures are described in a later chapter on exotic options.

Option pricing: setting the scene

The price of an option is a function of six different factors, which are the:

- strike price of the option
- current price of the underlying
- time to expiry
- risk-free rate of interest that applies to the life of the option
- volatility of the underlying asset's price returns
- value of any dividends or cash flows paid by the underlying asset during the life of the option.

We review the basic parameters next.

Pricing inputs

Let us consider the parameters of option pricing. Possibly the two most important are the current price of the underlying and the strike price of the option. The intrinsic value of a call option is the amount by which the strike price is below the price of the underlying, as this is the payoff if the option is exercised. Therefore the value of the call option will increase as the price of the underlying increases, and will fall as the underlying price falls. The value of a call will also decrease as the strike price increases. All this is reversed for a put option.

Generally for bond options a higher time to maturity results in higher option value. All other parameters being equal, a longer-dated option will always be worth at least as much as one that had a shorter life. Intuitively we would expect this because the holder of a longer-dated option has the same benefits as someone holding a shorter-dated option, in addition to a longer time period in which the intrinsic may increase. This rule is always true for American options, and usually true for European options. However certain factors, such as the payment of a coupon during the option life, may cause a longer-dated option to have only a slightly higher value than a shorter-date option.

The risk-free interest-rate is the rate applicable to the period of the option's life, so for our table of gilt options in the previous section, the option value reflected the three-month rate. The most common rate used is the T-bill rate, although for bond options it is more common to see the government bond repo rate being used. A rise in interest rates will increase the value of a call option, although not always for bond options. A rise in rates lowers the price of a bond, because it decreases the present value of future cash flows. However in the equity markets it is viewed as a sign that share price growth rates will increase. Generally however the relationship is the same for bond options as for equity options. The effect of a rise in interest rates for put options is the reverse: it causes the value to drop.

A coupon payment made by the underlying during the life of the option will reduce the price of the underlying asset on the ex-dividend date. This will result in a fall in the price of a call option and a rise in the price of a put option.

Bounds in option pricing

The upper and lower limits on the price of an option are relatively straightforward to set because prices must follow the rule of no-arbitrage pricing. A call option grants the buyer the right to buy a specified quantity of the underlying asset, at the level of the strike price, so therefore it is clear that the option could not have a higher value than the underlying asset itself. Therefore the upper limit or bound to the price of a call option is the price of the underlying asset. Therefore:

$$C \leq S$$

where C is the price of a call option and S is the current price of the underlying asset.

A put option grants the buyer the right to sell a specified unit of the underlying at the strike price X, therefore the option can never have a value greater than the strike price X. So we may set:

$$P \leq X$$

where P is the price of the put option. This rule will still apply for a European put option on its expiry date, so therefore we may further set that the option cannot have a value greater than the present value of the strike price X on expiry. That is,

$$P \leq X\mathrm{e}^{-rT}$$

where r is the risk-free interest for the term of the option and T is the maturity of the option in years.

The minimum limit or bound for an option is set according to whether the underlying asset is a dividend-paying security or not. For a call option written on a non-dividend paying security the lower bound is given by:

$$C \geq S - X\mathrm{e}^{-rT}$$

In fact as we noted early in this chapter a call option can only ever expire worthless, so its intrinsic value can never be worth less than zero. Therefore C > 0 and we then set the following:

$$C \geq \max[S - X\mathrm{e}^{-rT},0]$$

This reflects the law of no-arbitrage pricing. For put options on a non-dividend paying stock the lower limit is given by:

$$P \geq X\mathrm{e}^{-\mathrm{rT}} - S$$

and again the value is never less than zero so we may set:

$$P \geq \max[X\mathrm{e}^{-\mathrm{rT}} - S,0]$$

As we noted above, payment of a dividend by the underlying asset affects the price of the option. In the case of dividend paying stocks the upper and lower bounds for options are as follows:

$$C \geq S - D - X\mathrm{e}^{-\mathrm{rT}}$$

and

$$P \geq D + X\mathrm{e}^{-\mathrm{rT}} - S$$

where D is the present value of the dividend payment made by the underlying asset during the life of the option.

We can now look at option pricing in the Black–Scholes model, and this is considered in the next chapter.

Selected bibliography and references

Chance, D. *An Introduction to Options and Futures Markets*, Dryden Press, 1989.
Grabbe, J. O. *International Financial Markets*, 2nd edn, Elsevier, 1991.
Hull, J. *Options, Futures and Other Derivatives*, 4th edn, Prentice Hall, 1999.
Jarrow, R. and Turnbull, S. *Derivative Securities*, 2nd edn, South-Western, 1999.
Levy, H. *Introduction to Investments*, 2nd edn, South-Western Publishing, 1999, chs 22–25.
London International Financial Futures and Options Exchange, *LIFFE Options*, LIFFE, 1999.
Livingstone, M. *Money and Capital Markets*, 2nd edn, New York Institute of Finance, 1993, ch. 19.
McMillan, L. *Options as a Strategic Investment*, New York Institute of Finance, 1986.

18
Options II

In this chapter we present an overview of option pricing. There is a vast literature in this field, and space constraints allow us to consider only the basic concepts. Readers are directed to the bibliography for further recommended texts.

Option pricing

Previous interest rate products described in this book so far, both cash and derivatives, can be priced using rigid mathematical principles, because on maturity of the instrument there is a defined procedure that takes place such that one is able to calculate a fair value. This does not apply to options because there is uncertainty as to what the outcome will be on expiry; an option seller does not know whether the option will be exercised or not. This factor makes options more difficult to price than other financial market instruments. In this section we review the parameters used in the pricing of an option, and introduce the Black–Scholes pricing model.

Pricing an option is a function of the probability that it will be exercised. Essentially the premium paid for an option represents the buyer's *expected profit* on the option. Therefore, as with an insurance premium, the writer of an option will base his price on the assessment that the payout on the option will be equal to the premium, and this is a function on the probability that the option will be exercised. Option pricing therefore bases its calculation on the assessment of the probability of exercise, and derives from this an expected outcome, and hence a fair value for the option premium. The expected payout, as with an insurance company premium, should equal the premium received.

The following factors influence the price of an option.

- *The behaviour of financial prices*. One of the key assumptions made by the Black–Scholes model (B–S) is that asset prices follow a lognormal distribution. Although this is not strictly accurate, it is close enough of an approximation to allow its use in option pricing. In fact observation shows that while prices themselves are not normally distributed, asset returns are, and we define returns as

$$\ln\left(\frac{P_{t+1}}{P_t}\right)$$

where P_t is the market price at time t and P_{t+1} is the price one period later. The distribution of prices is called a lognormal distribution because the logarithm of the prices is normally distributed; the asset returns are defined as the logarithm of the price relatives and are assumed to follow the normal distribution. The expected return as a result of assuming this distribution is given by

$$\mathrm{E}\left[\ln\left(\frac{P_t}{P_0}\right)\right] = rt$$

where $E[\ ]$ is the expectation operator and r is the annual rate of return. The derivation of this expression is given in Appendix 18.1.

- *The strike price*. The difference between the strike price and the underlying price of the asset at the time the option is struck will influence the size of the premium, as this will impact on the probability that the option will be exercised. An option that is deeply in-the-money has a greater probability of being exercised.
- *Volatility*. The volatility of the underlying asset will influence the probability that an option is exercised, as a higher volatility indicates a higher probability of exercise. This is considered in detail below.
- *The term to maturity*. A longer-dated option has greater time value and a greater probability of eventually being exercised.
- *The level of interest rates*. The premium paid for an option in theory represents the expected gain to the buyer at the time the option is exercised. It is paid up-front so it is discounted to obtain a present value. The discount rate used therefore has an effect on the premium, although it is less influential than the other factors presented here.

The *volatility* of an asset measures the variability of its price returns. It is defined as the annualised standard deviation of returns, where variability refers to the variability of the returns that generate the asset's prices, rather than the prices directly. The standard deviation of returns is given by (18.1):

$$\sigma = \sqrt{\sum_{i=1}^{N} \frac{(x_i - \mu)^2}{N-1}} \tag{18.1}$$

where x_i is the ith price relative, μ the arithmetic mean of the observations and N the total number of observations. The value is converted to an annualised figure by multiplying it by the square root of the number of days in a year, usually taken to be 250 working days. Using this formula from market observations it is possible to calculate the *historic volatility* of an asset. The volatility of an asset is one of the inputs to the B–S model.

Of the inputs to the B–S model, the variability of the underlying asset, or its volatility, is the most problematic. The distribution of asset prices is assumed to follow a lognormal distribution, because the logarithm of the prices is normally distributed (we assume lognormal rather than normal distribution to allow for the

fact that prices cannot – as could be the case in a normal distribution – have negative values): the range of possible prices starts at zero.

Note that it is the asset price *returns* on which the standard deviation is calculated, and not the actual prices themselves. This is because using prices would produce inconsistent results, as the actual standard deviation itself would change as price levels increased.

However calculating volatility using the standard statistical method gives us a figure for *historic* volatility. What is required is a figure for *future* volatility, since this is relevant for pricing an option expiring in the future. Future volatility cannot be measured directly, by definition. Market makers get around this by using an option pricing model 'backwards'. An option pricing model calculates the option price from volatility and other parameters. Used in reverse the model can calculate the volatility implied by the option price. Volatility measured in this way is called *implied volatility*. Evaluating implied volatility is straightforward using this method, and generally more appropriate than using historic volatility, as it provides a clearer measure of an option's fair value. Implied volatilities of deeply in-the-money or out-of-the-money options tend to be relatively high.

The Black–Scholes option model

Most option pricing models are based on one of two methodologies, although both types employ essentially identical assumptions. The first method is based on the resolution of the partial differentiation equation of the asset price model, corresponding to the expected payoff of the option security. This is the foundation of the B–S model. The second type of model uses the martingale method, and was first introduced by Harrison and Kreps (1979) and Harrison and Pliska (1981), where the price of an asset at time 0 is given by its discounted expected future payoffs, under the appropriate probability measure, known as the *risk-neutral probability*. There is a third type of model that assumes lognormal distribution of asset returns but follows the two-step binomial process.

In order to employ the pricing models, we accept a state of the market that is known as a *complete market*,[1] one where there is a viable financial market. This is where the rule of no-arbitrage pricing exists, so that there is no opportunity to generate risk-free arbitrage due to the presence of, say, incorrect forward interest rates. The fact that there is no opportunity to generate risk-free arbitrage gains means that a zero-cost investment strategy that is initiated at time t will have a zero maturity value. The martingale property of the behaviour of asset prices states that an accurate estimate of the future price of an asset may be obtained from current price information. Therefore the relevant information used to calculate forward asset prices is the latest price information. This was also a property of the semi-strong and strong-form market efficiency scenarios described by Fama (1965).

In this section we describe the B–S option model in accessible fashion; more technical treatments are given in the relevant references listed in the bibliography.

[1] First proposed by Arrow and Debreu (1953, 1954).

Assumptions

The B–S model describes a process to calculate the fair value of a European call option under certain assumptions, and apart from the price of the underlying asset S and the time t all the variables in the model are assumed to be constant, including most crucially the volatility. The following assumptions are made:

- There are no transaction costs, and the market allows short selling.
- Trading is continuous.
- Underlying asset prices follow geometric Brownian motion, with the variance rate proportional to the square root of the asset price.
- The asset is a non-dividend paying security.
- The interest rate during the life of the option is known and constant.
- The option can only be exercised on expiry.

The B–S model is neat and intuitively straightforward to explain, and one of its many attractions is that it can readily be modified to handle other types of options such as foreign exchange or interest-rate options. The assumption of the behaviour of the underlying asset price over time is described by (18.2), which is a generalised Weiner process, and where a is the expected return on the underlying asset and b is the standard deviation of its price returns.

$$\frac{\mathrm{d}S}{S} = a\mathrm{d}t + b\mathrm{d}W \tag{18.2}$$

The B–S model and pricing derivative instruments

We assume a financial asset is specified by its terminal payoff value, therefore when pricing an option we require the fair value of the option at the initial time when the option is struck, and this value is a function of the expected terminal payoff of the option, discounted to the day when the option is struck. In this section we present an intuitive explanation of the B–S model, in terms of the normal distribution of asset price returns.

From the definition of a call option, we can set the expected value of the option at maturity T as:

$$\mathrm{E}(C_T) = E[\max(S_T - X, 0)] \tag{18.3}$$

where

S_T is the price of the underlying asset at maturity T
X is the strike price of the option.

From (18.3) we know that there are only two possible outcomes that can arise on maturity, either the option will expire in-the-money and the outcome is $S_T - X$, or the option will be out-of-the-money and the outcome will be 0. If we set the term

p as the probability that on expiry $S_T > X$, equation (18.3) can be rewritten as (18.4).

$$E(C_T) = \rho \times (E[S_T \mid S_T > X] - X) \tag{18.4}$$

where $E[S_T \mid S_T > X]$ is the expected value of S_T given that $S_T > X$. Equation (18.4) gives us an expression for the expected value of a call option on maturity. Therefore to obtain the fair price of the option at the time it is struck, the value given by (18.4) must be discounted back to its present value, and this is shown as (18.5).

$$C = p \times e^{-rt} \times (E[S_T \mid S_T > X] - X) \tag{18.5}$$

where r is the continuously compounded risk-free rate of interest, and t is the time from today until maturity. Therefore to price an option we require the probability p that the option expires in-the-money, and we require the expected value of the option given that it does expire in-the-money, which is the last term of (18.5). To calculate p we assume that asset prices follow a stochastic process, which enables us to model the probability function.

The B-S model is based on the resolution of the following partial differential equation:

$$\tfrac{1}{2}\sigma^2 S^2 \left(\frac{\partial^2 C}{\partial S^2}\right) + rS\left(\frac{\partial C}{\partial S}\right) + \left(\frac{\partial C}{\partial t}\right) - rC = 0 \tag{18.6}$$

under the appropriate parameters. We do not demonstrate the process by which this equation is arrived at. The parameters refer to the payoff conditions corresponding to a European call option, which we considered above. We do not present a solution to the differential equation at (18.6), which is beyond the scope of this book, but we can consider now how the probability and expected value functions can be solved. For a fuller treatment readers may wish to refer to the original account by Black and Scholes; other good accounts are given in Ingersoll (1987), Neftci (1996), and Nielsen (1999) among others.

The probability p that the underlying asset price at maturity exceeds X is equal to the probability that the return over the time period the option is held will exceed a certain critical value. Remember that we assume normal distribution of asset price returns. As asset returns are defined as the logarithm of price relatives, we require p such that:

$$p = prob[S_T > X] = prob\left[return > \ln\left(\frac{X}{S_0}\right)\right] \tag{18.7}$$

where S_0 is the price of the underlying asset at the time the option is struck. Generally the probability that a normal distributed variable x will exceed a critical value x_c is given by (18.8):

$$P[x > x_c] = 1 - N\left(\frac{x_c - \mu}{\sigma}\right) \tag{18.8}$$

where μ and σ are the mean and standard deviation of x respectively and $N(\ )$ is the cumulative normal distribution. We know from our earlier discussion of the behaviour of asset prices that an expression for μ is the natural logarithm of the asset price returns; we already know that the standard deviation of returns is $\sigma\sqrt{t}$. Therefore with these assumptions, we may combine (18.7) and (18.8) to give us (18.9), that is,

$$prob[S_T > \mathrm{X}] = prob\left[return > \ln\left(\frac{X}{S_0}\right)\right] = 1 - N\left(\frac{\ln\left(\frac{X}{S_0}\right) - \left(r - \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}}\right) \tag{18.9}$$

Under the conditions of the normal distribution, the symmetrical shape means that we can obtain the probability of an occurrence based on $1 - N(d)$ being equal to $N(-d)$. Therefore we are able to set the following relationship, at (18.10):

$$p = prob[S_T > X] = N\left(\frac{\ln\left(\frac{S_0}{X}\right) + \left(r - \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}}\right) \tag{18.10}$$

Now we require a formula to calculate the expected value of the option on expiry, the second part of the expression at (18.5). This involves the integration of the normal distribution curve over the range from X to infinity. This is not shown here; however the result is given at (18.11).

$$E[S_T \mid S_T > X] = S_0 e^{rt} \frac{N(d_1)}{N(d_2)} \tag{18.11}$$

where

$$d_1 = \left(\frac{\ln\left(\frac{S_0}{X}\right) + \left(r - \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}}\right)$$

and

$$d_2 = \left(\frac{\ln\left(\frac{S_0}{X}\right) + \left(r - \frac{\sigma^2}{2}\right)t}{\sigma\sqrt{t}}\right) = d_1 - \sigma\sqrt{t}$$

We now have expressions for the probability that an option expires in-the-money as well as the expected value of the option on expiry, and we incorporate these

into the expression at (18.5), which gives us (18.12).

$$C = N(d_2) \times e^{-rt} \times \left(S_0 e^{-rt} \frac{N(d_1)}{N(d_2)} - X \right) \tag{18.12}$$

Equation (18.12) can be rearranged to give (18.13), which is the well-known Black–Scholes option pricing model for a European call option.

$$C = S_0 N(d_1) - X e^{-rt} N(d_1) \tag{18.13}$$

where

S_0 is the price of the underlying asset at the time the option is struck
X is the strike price
r is the continuously compounded risk-free interest rate
t is the maturity of the option.

What the expression at (18.13) states is that the fair value of a call option is the expected present value of the option on its expiry date, assuming that prices follow a lognormal distribution.

$N(d_1)$ and $N(d_2)$ are the cumulative probabilities from the normal distribution of obtaining the values d_1 and d_2, given above. $N(d_1)$ is the *delta* of the option. The term $N(d_2)$ represents the probability that the option will be exercised. The term e^{-rt} is the present value of one unit of cash received t periods from the time the option is struck. Where $N(d_1)$ and $N(d_2)$ are equal to 1, which is the equivalent of assuming complete certainty, the model is reduced to:

$$C = S - X\mathrm{e}^{-rt}$$

which is the expression for Merton's lower bound for continuously compounded interest rates, and which we introduced in intuitive fashion in Chapter 17. Therefore under complete certainty the B–S model reduces to Merton's bound.

The put–call parity relationship

Up to now we have concentrated on calculating the price of a call option. However the previous section introduced the boundary condition for a put option, so it should be apparent that this can be solved as well. In fact the price of a call option and a put option are related via what is known as the *put–call parity theorem*. This is an important relationship, and obviates the need to develop a separate model for put options.

Consider a portfolio Y that consists of a call option with a maturity date T and a zero-coupon bond that pays X on the expiry date of the option. Consider also a second portfolio Z that consists of a put option also with maturity date T and one share. The value of portfolio Y on the expiry date is given by (18.14):

$$MV_{Y,T} = \max[S_T - X, 0] + X = \max[X, S_T] \tag{18.14}$$

The value of the second portfolio Z on the expiry date is:

$$MV_{Z,T} = \max[X - S_T, 0] + S_T = \max[X, S_T] \tag{18.15}$$

Both portfolios have the same value at maturity. Therefore they must also have the same initial value at start time t, otherwise there would be an arbitrage opportunity. Prices must be arbitrage-free, therefore the following put–call relationship must hold:

$$C_t - P_t = St - Xe^{-r(T-t)} \tag{18.16}$$

If the relationship at (18.16) did not hold, then arbitrage would be possible. So using this relationship, the value of a European put option is given by the B–S model as shown below, at (18.17).

$$P(S,T) = -SN(-d_1) + Xe^{-rT}N(-d_2) \tag{18.17}$$

Example 18.1: The Black–Scholes model

Here we illustrate a simple application of the B–S model. Consider an underlying asset, usually assumed to be a non-dividend paying equity, with a current price of 25, and volatility of 23%. The short-term risk-free interest rate is 5%. An option is written with strike price 21 and a maturity of three months. Therefore we have:

$S = 25$
$X = 21$
$r = 5\%$
$T = 0.25$
$\sigma = 23\%$

To calculate the price of the option, we first calculate the discounted value of the strike price, as follows:

$$Xe^{-rT} = 21e^{-0.05(0.25)} = 20.73913$$

We then calculate the values of d_1 and d_2:

$$d_1 = \frac{\ln(25/21) + |0.05 + (0.5 + (0.5)(0.23)^2|0.25}{0.23\sqrt{0.25}} = \frac{0.193466}{0.115} = 1.682313$$

$$d_2 = d_1 - 0.23\sqrt{0.25} = 1.567313$$

We now insert these values into the main price equation:

$$C = 25N(1.682313) - 21e^{-0.05(0.25)}N(1.567313)$$

Using the approximation of the cumulative normal distribution at the points 1.68 and 1.56, the price of the call option is:

$$C = 25(0.9535) - 20.73913(0.9406) = 4.3303$$

What would be the price of a put option on the same stock? The values of $N(d_1)$ and $N(d_2)$ are 0.9535 and 0.9406, therefore the put price is calculated as:

$$P = 20.7391\ (1 - 0.9406) - 25\ (1 - 0.9535) = 0.06943$$

If we use the call price and apply the put–call parity theorem, the price of the put option is given by:

$$P = C - S + Xe^{-rT} = 4.3303 - 25 + 21e^{-0.05(0.25)} = 0.069434$$

This is exactly the same price that was obtained by the application of the put option formula in the B–S model above.

As we noted early in this chapter, the premium payable for an option will increase if the time to expiry, the volatility or the interest rate is increased (or any combination is increased). Thus if we maintain all the parameters constant but price a call option that has a maturity of six months or $T = 0.5$, we obtain the following values:

$d_1 = 1.3071$, giving $N(d_1) = 0.9049$
$d_2 = 1.1445$, giving $N(d_2) = 0.8740$

The call price for the longer-dated option is 4.7217.

The Black-Scholes model as an Excel spreadsheet

In Appendix 18.3 we show the spreadsheet formulae required to build the B–S model into Microsoft® Excel. The user must ensure that the Analysis Tool-Pak add-in is available, otherwise some of the function references may not work. By setting up the cells in the way shown, the fair value of a vanilla call or put option may be calculated. The put–call parity is used to enable calculation of the put price.

Black–Scholes and the valuation of bond options

In this section we illustrate the application of the B–S model to the pricing of an option on a zero-coupon bond and a plain vanilla fixed-coupon bond.

For a zero-coupon bond the theoretical price of a call option written on the bond is given by (18.18):

$$C = PN(d_1) - Xe^{-rT}N(d_2) \tag{18.18}$$

where P is the price of the underlying bond and all other parameters remain the same. If the option is written on a coupon-paying bond, it is necessary to subtract the present value of all coupons paid during the life of the option from the bond's price. Coupons sometimes lower the price of a call option because a coupon makes it more attractive to hold a bond rather than an option on the bond. Call options on bonds are often priced at a lower level than similar options on zero-coupon bonds.

Example 18.2: The B–S model and bond option pricing

Consider a European call option written on a bond that has the following characteristics:

Price	£98
Coupon	8.00% (semi-annual)
Time to maturity	5 years
Bond price volatility	6.02%
Coupon payments	£4 in three months and nine months
Three-month interest rate	5.60%
Nine-month interest rate	5.75%
One-year interest rate	6.25%

The option is written with a strike price of £100 and has a maturity of one year. The present value of the coupon payments made during the life of the option is £7.78, as shown below.

$$4e^{-0.056\times0.25} + 4e^{-0.0575\times0.75} = 3.9444 + 3.83117 = 7.77557$$

This gives us $P = 98 - 7.78 = £90.22$

Applying the B–S model we obtain:

$$d_1 = [\ln(90.22/100) + 0.0625 + 0.001812]/0.0602 = -0.6413$$
$$d_2 = d_1 - (0.0602 \times 1) = 0.7015$$
$$C = 90.22N(-0.6413) - 100e^{-0.0625}N(-0.7015) = 1.1514$$

Therefore the call option has a value of £1.15, which will be composed entirely of time value. Note also that a key assumption of the model is constant interest rates, yet it is being applied to a bond price – which is essentially an interest rate – that is considered to follow stochastic price processes.

Interest-rate options and the Black model

In 1976 Fisher Black presented a slightly modified version of the B–S model, using similar assumptions, to be used in pricing forward contracts and interest-rate options. The Black model is used in banks today to price instruments such as swaptions in addition to bond and interest-rate options like caps and floors.

In this model the spot price $S(t)$ of an asset or a commodity is the price payable for immediate delivery today (in practice, up to two days forward) at time t. This price is assumed to follow a geometric Brownian motion. The theoretical price for a futures contract on the asset, $F(t,T)$ is defined as the price agreed today for delivery of the asset at time T, with the price agreed today but payable on delivery. When $t = T$, the futures price is equal to the spot price. A futures contract is cash settled every day via the clearing mechanism, whereas a forward contract is a contract to buy or sell the asset where there is no daily mark-to-market and no daily cash settlement.

Let us set f as the value of a forward contract, u as the value of a futures contract and C as the value of an option contract. Each of these contracts is a function of the futures price $F(t,T)$, as well as additional variables. So we may write at time t the values of all three contracts as $f(F,t)$, $u(F,t)$ and $C(F,t)$. The value of the forward contract is also a function of the price of the underlying asset S at time T and can be written $f(F,t,S,T)$. Note that the value of the forward contract f is not the same as the price of the forward contract. The forward price at any given time is the delivery price that would result in the contract having a zero value. At the time the contract is transacted, the forward value is zero. Over time both the price and the value will fluctuate. The futures price, on the other hand, is the price at which a forward contract has a zero current value. Therefore at the time of the trade the forward price is equal to the futures price F, which may be written as:

$$f(F,tF,T) = 0 \tag{18.19}$$

Equation (18.19) simply states that the value of the forward contract is zero when the contract is taken out, and the contract price S is always equal to the current futures price, $F(t,T)$.[2]

The principal difference between a futures contract and a forward contract is that a futures contract may be used to imply the price of forward contracts. This arises from the fact that futures contracts are repriced each day, with a new contract price that is equal to the new futures price. Hence when F rises, such that $F>S$, the forward contract has a positive value, and when F falls, the forward contract has a negative value. When the transaction expires and delivery takes place, the futures price is equal to the spot price and the value of the forward contract is equal to the spot price minus the contract price or the spot price.

$$f(F,T,S,T) = F - S \tag{18.20}$$

On maturity the value of a bond or commodity option is given by the maximum of zero, and the difference between the spot price and the contract price. Since at that date the futures price is equal to the spot price, we conclude that:

$$C(F,T) = \begin{cases} F - S & \text{if } F \geq S \\ 0 & \text{else} \end{cases} \tag{18.21}$$

[2] This assumption is held in the market but does not hold good over long periods, due chiefly to the difference in the way futures and forwards are marked-to-market, and because futures are cash settled on a daily basis while forwards are not.

The assumptions made in the Black model are that the prices of futures contracts follow a lognormal distribution with a constant variance, and that the Capital Asset Pricing Model applies in the market. There is also an assumption of no transaction costs or taxes. Under these assumptions, we can create a risk-free hedged position that is composed of a long position in the option and a short position in the futures contract. Following the B–S model, the number of options put on against one futures contract is given by $[\partial C(F,t)/\partial F]$, which is the derivative of $C(F, t)$ with respect to F. The change in the hedged position resulting from a change in price of the underlying is given by (18.22):

$$\partial C(F,t) - [\partial C(F,t)/\partial F]\partial F \tag{18.22}$$

Due to the principle of arbitrage-free pricing, the return generated by the hedged portfolio must be equal to the risk-free interest rate, and this together with an expansion produces the following partial differential equation:

$$\left[\frac{\partial \mathrm{C}(F,t)}{\partial t}\right] = rC(F,t) - \tfrac{1}{2}\sigma^2 F^2 \left[\frac{\partial^2 C(F,t)}{\partial F^2}\right] \tag{18.23}$$

which is solved by setting the following:

$$\tfrac{1}{2}\sigma^2 F^2 \left[\frac{\partial^2 \mathrm{C}(F,t)}{\partial F^2}\right] - rC(F,t) + \left[\frac{\partial C(F,t)}{\partial t}\right] = 0 \tag{18.24}$$

The solution to the partial differential equation (18.23) is not presented here.

The result, by denoting $T = t - T$ and using (18.23), gives the fair value of a commodity option or option on a forward contract as shown below, at (18.25).

$$C(F,t) = \mathrm{e}^{-\mathrm{rT}}[FN(d_1) - S_T N(d_2)] \tag{18.25}$$

where

$$d_1 = \frac{1}{\sigma\sqrt{T}}\left[\ln\left(\frac{F}{ST}\right) + (\tfrac{1}{2}\sigma^2)T\right]$$

$$d_2 = d_1 - \sigma\sqrt{T}$$

There are a number of other models that have been developed for specific contracts, for example the Garman and Kohlhagen (1983) and Grabbe (1983) models, used for currency options, and the Merton, Barone-Adesi and Whaley or BAW model (1987) used for commodity options. For the valuation of American options, on dividend-paying assets, another model has been developed by Roll, Geske and Whaley. More recently the Black–Derman–Toy model (1990) has been used to price exotic options. A detailed discussion of these, though very interesting, is outside the scope of this book.

Comment on the Black–Scholes model

The introduction of the B–S model was one of the great milestones in the development of the global capital markets, and it remains an important pricing model today. Many of the models introduced later for application to specific products are still based essentially on the B–S model. Subsequently academics have presented some weaknesses in the model that stem from the nature of the main assumptions behind the model itself, which we will summarise here. The main critique of the B–S model appears to centre on:

- Assumption of frictionless markets; this is at best only approximately true for large market counterparties.
- Constant interest rate: this is possibly the most unrealistic assumption. Interest rates over even the shortest time frame (the overnight rate) fluctuate considerably. In addition to a dynamic short rate, the short-end of the yield curve often moves in the opposite direction to moves in underlying asset prices, particularly so with bonds and bond options.
- Lognormal distribution; this is accepted by the market as a reasonable approximation but not completely accurate, and also misses out most extreme moves or market shocks.
- European option only; although it is rare for American options to be exercised early, there are situations when it is optimal to do so, and the B–S model does not price these situations.
- For stock options, the assumption of a continuous constant dividend yield is clearly not realistic, although the trend in the US markets is for ordinary shares to cease paying dividends altogether.

These points notwithstanding, the B–S model paved the way for the rapid development of options as liquid tradeable products and is widely used today.

Stochastic volatility

The B–S model assumes a constant volatility, and for this reason, and because it is based on mathematics, often fails to pick up on market 'sentiment' when there is a large downward move or shock. This is not a failing limited to the B–S model. For this reason, however, it undervalues out-of-the-money options, and to compensate for this market makers push up the price of deep in or out-of-the-money options, giving rise to the *volatility smile*. This is considered in the next chapter.

The effect of stochastic volatility not being catered for then is to introduce mispricing, specifically the undervaluation of out-of-the-money options and the overvaluation of deeply in-the-money options. This is because when the price of the underlying asset rises, its volatility level also increases. The effect of this is that assets priced at relatively high levels do not tend to follow the process described by geometric Brownian motion. The same is true for relatively low asset prices and price volatility, but in the opposite direction. To compensate for this stochastic volatility models have been developed, such as the Hull–White model (1987).

Implied volatility

The volatility parameter in the B–S model, by definition, cannot be observed directly in the market as it refers to volatility going forward. It is different from historic volatility which can be measured directly, and this value is sometimes used to estimate implied volatility of an asset price. Banks therefore use the value for *implied volatility*, which is the volatility obtained using the prices of exchange-traded options. Given the price of an option and all the other parameters, it is possible to use the price of the option to determine the volatility of the underlying asset implied by the option price. The B–S model however cannot be rearranged into a form that expresses the volatility measure as a function of the other parameters. Generally therefore a numerical iteration process is used to arrive at the value for given the price of the option, usually the Newton–Raphson method.

The market uses implied volatilities to gauge the volatility of individual assets relative to the market. Volatility levels are not constant, and fluctuate with the overall level of the market, as well as for stock-specific factors. When assessing volatilities with reference to exchange-traded options, market makers will use more than one value, because an asset will have different implied volatilities depending on how in-the-money the option itself is. The price of an at-the-money option will exhibit greater sensitivity to volatility than the price of a deeply in or out-of-the-money option. Therefore market makers will take a combination of volatility values when assessing the volatility of a particular asset.

A final word on option models

We have only discussed the B–S model and the Black model in this chapter. Other pricing models have been developed that follow on from the pioneering work done by Black and Scholes. The B–S model is essentially the most straightforward and the easiest to apply, and subsequent research has focused on easing some of the restrictions of the model in order to expand its applicability. Areas that have been focused on include a relaxation of the assumption of constant volatility levels for asset prices, as well as work on allowing for the valuation of American options and options on dividend-paying stocks. However often in practice some of the newer models require input of parameters that are difficult to observe or measure directly, which limits their application as well.

Often there is a difficulty in calibrating a model due to the lack of observable data in the marketplace. The issue of calibration is an important one in the implementation of a pricing model, and involves inputting actual market data and using this as the parameters for calculation of prices. So for instance a model used to calculate the prices of sterling market options would use data from the UK market, including money market, futures and swaps rates to build the zero-coupon yield curve, and volatility levels for the underlying asset or interest rate (if it is a valuing options on interest-rate products, such as caps and floors). What sort of volatility is used? In some banks actual historical volatilities are used, more usually volatilities implied by exchange-traded option prices. Another crucial piece of data

for multi-factor models (following Heath–Jarrow–Morton and other models based on this) is the correlation coefficients between forward rates on the term structure. This is used to calculate volatilities using the model itself.

The issue of calibrating the model is important, because incorrectly calibrated models will produce errors in option valuation. This can have disastrous results, which may be discovered only after significant losses have been suffered. If data is not available to calibrate the model, it may be that a simpler one needs to be used. The lack of data is not an issue for products priced in, say, dollars, sterling or euro, but may be in other currency products if data is not so readily available. This might explain why the B–S model is still widely used today, although markets observe an increasing use of models such as the Black–Derman–Toy (1990) and Brace–Gatarek–Musiela (1994) for more exotic option products.

Many models, because of the way that they describe the price process, are described as Gaussian interest rate models. The basic process is described by an Itô process:

$$\mathrm{d}P_T / P_T = \mu_{\underline{T}}\mathrm{d} + \sigma_T \mathrm{d}W$$

where P_T is the price of a zero-coupon bond with maturity date T, and W is a standard Weiner process. The basic statement made by Gaussian interest-rate models is that:

$$P_T(t) = E_T \exp(-\int_{\mathrm{t}}^{T} r(s)\mathrm{d}s)$$

Models that capture the process in this way include Cox–Ingersoll–Ross and Harrison and Pliska. We are summarising here only, but essentially such models state that the price of an option is equal to the discounted return from a risk-free instrument. This is why the basic B–S model describes a portfolio of a call option on the underlying stock and a cash deposit invested at the risk-free interest rate. This was reviewed in the chapter. We then discussed how the representation of asset prices as an expectation of a discounted payoff from a risk-free deposit does not capture the real-world scenario presented by many option products. Hence the continuing research into developments of the basic model.

Following on from B–S, under the assumption that the short-term spot rate drives bond and option prices, the basic model can be used to model an interest-rate term structure, as given by Vasicek and Cox–Ingersoll–Ross. The short-term spot rate is assumed to follow a diffusion process

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

which is a standard Weiner process. From this it is possible to model the complete term structure based on the short-term spot rate and the volatility of the short-term rate. This approach is modified by Heath–Jarrow–Morton (1992), which was reviewed earlier as an interest-rate model.

Appendix 18.1: Summary of basic statistical concepts

The *arithmetic mean* μ is the average of a series of numbers. The *variance* is the sum of the squares of the difference of each observation from the mean, and from the variance we obtain the *standard deviation* σ which is the square root of the variance. The *probability density* of a series of numbers is the term for how likely any of them is to occur. In a normal probability density function, described by the normal distribution, the probability density is given by:

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

Most option pricing formulas assume a normal probability density function, specifically that movements in the natural logarithm of asset prices follow this function. That is,

$$\ln\left(\frac{\text{today's price}}{\text{yesterday's price}}\right)$$

is assumed to follow a normal probability density function. This relative price change is equal to:

$$\left(1+r\times\frac{days}{year}\right)$$

where r is the rate of return being earned on an investment in the asset. The value

$$\ln\left(1+r\times\frac{days}{year}\right)$$

is equal to $r\times\frac{days}{year}$ where r is the continuously compounded rate of return. Therefore the value

$$\ln\left(\frac{\text{today's price}}{\text{yesterday's price}}\right)$$

is equal to the continuously compounded rate of return on the asset over a specified holding period.

Appendix 18.2: Lognormal distribution of returns

In the distribution of asset price returns, returns are defined as the logarithm of price relatives and are assumed to follow the normal distribution, given by:

$$\ln\left(\frac{P_t}{P_0}\right)\sim N(rt,\sigma\sqrt{t}) \tag{18.26}$$

where

P_t	is the price at time t
P_0	is the price at time 0
$N(m,s)$	is a random normal distribution with mean m and standard deviation s
r	is the annual rate of return
σ	is the annualised standard deviation of returns.

From (18.26) we conclude that the logarithm of the prices is normally distributed, due to (18.27) where P_0 is a constant:

$$\ln(P_t) \sim \ln(P_0) + N(rt, \sigma\sqrt{t}) \tag{18.27}$$

We conclude that prices are normally distributed and are described by the relationship,

$$\frac{P_t}{P_0} \sim e^{N(rt,\sigma\sqrt{t})}$$

and from this relationship we may set the expected return as rt.

Appendix 18.3: The Black–Scholes model in Microsoft® Excel

To value a vanilla option under the following parameters, we can use Microsoft Excel to carry out the calculation as shown in Table 18.1.

Price of underlying	100
Volatility	0.0691
Maturity of option	3 months
Strike price	99.5
Risk-free rate	5%

Table 18.1 Microsoft Excel calculation of vanilla option price

Cell	*C*	*D*	
8	Underlying price, S	100	
9	Volatility %	0.0691	
10	Option maturity years	0.25	
11	Strike price, X	99.50	
12	Risk-free interest rate %	0.05	
13			
14			
15			*Cell formulae:*
16	ln (S/X)	0.005012542	=LN (D8/D11)
17	Adjusted return	0.0000456012500	=((D12-D9)^2/ 2)*D10
18	Time adjusted volatility	0.131434394	=(D9*D10)^0.5
19	d2	0.038484166	=(D16+D17)/D18

Continued

Table 18.1 Continued

Cell	*C*	*D*	
20	N(d2)	0.515349233	=NORMSDIST(D19)
21			
22	d1	0.16991856	=D19+D18
23	N(d1)	0.56746291	=NORMSDIST(D22)
24	e-rt	0.9875778	=EXP(-D10*D12)
25			
26	CALL	6.106018498	=D8*D23-D11*D20*D24
27	PUT	4.370009648	* =D26-D8+D11*D24
*	By put-call parity, P = C − S + Xe-rt		

Selected bibliography and references

Arrow, K. and Debreu, G. 'Existence of an equilibrium for a competitive economy', *Econometrica*, 22, 1954, pp. 265–90.

Black, F., Derman, E. and Toy, W. 'A one-factor model of interest rates and its application to treasury bond options', in Hughston, L. (ed.), *Vasicek and Beyond*, Risk Publications, 1996.

Black, F. and Scholes, M. 'The pricing of options and corporate liabilities', *Journal of Political Economy*, 81, May–June 1973, pp. 637–59.

Bookstaber, R. *Option Pricing and Strategies in Investing*, Addison-Wesley, 1982.

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

Brace, A., Gatarek, D. and Musiela, M. 'The market model of interest rate dynamics', in Hughston, L. (ed.), *Vasicek and Beyond*, Risk Publications, 1996.

Briys, E., Bellalah, M. and Minh, M. H. *Options, Futures and Exotic Derivatives*, Wiley, 1998.

Choudhry, M. *The Bond and Money Markets: Strategy, Trading, Analysis*, Butterworth-Heinemann, 2001, chs 42–49.

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

Cox, J., Ingersoll, J. and Ross, S. 'An inter-temporal general equilibrium model of asset prices', *Econometrica*, 53, 1985.

Cox, J. and Ross, S. 'The valuation of options for alternative stochastic processes', *Journal of Financial Economics*, 3, 1976, pp. 145–66.

Cox, J., Ross, S. and Rubinstein, M. 'Option pricing: a simplified approach', *Journal of Financial Economics*, 7, Oct. 1979, pp. 229–64.

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

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

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

Harrison, J. and Pliska, S. 'Martingales and stochastic integrals in the theory of continuous trading', *Stochastic Processes and Their Applications*, 11, 1981, pp. 216–60.

Haug, E. G. *The Complete Guide to Option Pricing Formulas*, McGraw-Hill, 1998.

Heath, D., Jarrow, R. and Morton, A. 'Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation', *Econometrica*, 60, Jan. 1992, pp. 77–105.

Heston, S. 'A closed form solution for options with stochastic volatility with application to bond and currency options', *Review of Financial Studies*, 6, 1993, pp. 327–43.

Heston, S. 'Invisible parameters in option prices', *Journal of Finance*, 48, 3, 1993, pp. 933–47.
Hull, J. and White, A. 'The pricing of options on assets with stochastic volatilities', *Journal of Finance*, 42, 1987, pp. 281–300.
Hull, J. and White, A., 'An analysis of the bias caused by a stochastic volatility in option pricing', *Advances in Futures and Options Research*, 3, 1988, pp. 29–61.
Ingersoll, J. *Theory of Financial Decision Making*, Rowman & Littlefield, 1987.
Ito, K. 'On stochastic differential equations', *American Mathematical Society*, 4, 1951, pp. 1–51.
Klemkosky, R. and Resnick, B. 'Put–call parity and market efficiency', *Journal of Financial Economics*, 34, Dec.1979, pp. 1141–55.
Merton, R. 'Theory of rational option pricing', *Bell Journal of Economics and Management Science*, 4, Spring 1973, pp. 141–83.
Merton, R. 'Option pricing when underlying stock returns are discontinuous', *Journal of Financial Economics*, 3, Jan.–Mar. 1976, pp. 125–44.
Merton, R. *Continuous-Time Finance*, Blackwell, 1990.
Neftci, S. *An Introduction to the Mathematics of Financial Derivatives*, Academic Press, 1996.
Nielsen, L. T. *Pricing and Hedging of Derivative Securities*, Oxford UP, 1999.
Rendleman, R. and Bartter, B. 'Two state option pricing', *Journal of Finance*, 34, 1979, pp. 1092–110.
Rubinstein, M. 'Nonparametric tests of alternative option pricing models', *Journal of Financial Economics*, 40, 1985, pp. 455–80.
Scott, L. 'Option pricing when the variance changes randomly', *Journal of Financial and Quantitative Analysis*, 22, 1987, pp. 419–38.
Stein, E. and Stein, J. 'Stock price distributions with stochastic volatility: an analytic approach', *Review of Financial Studies*, 4, 1991, pp. 113–35.
Vasicek, O. 'An equilibrium characterization of the term structure', *Journal of Financial Economics*, 5, 1977, pp. 177–88.
Whaley, R. 'On the valuation of American call options on stocks with known dividends', *Journal of Financial Economics*, 9, 1981, pp. 207–11.

19
Options III

We continue with options in this chapter, with a look at how options behave in response to changes in market conditions. To start we consider the main issues that a market maker in options must consider when writing options. We then review 'the Greeks', the measures by which the sensitivity of an option book is calculated. We conclude with a discussion on an important set of interest-rate options in the market, *caps* and *floors*.

Behaviour of option prices

As we noted in the previous chapter, the value of an option is a function of five factors:

- the price of the underlying asset
- the strike price of the option
- the time to expiry of the option
- the volatility level of the underlying asset price returns
- the risk-free interest rate applicable to the life of the option.

The Black–Scholes (B–S) model assumes that the level of volatility and interest rates stays constant, so that changes in these will impact on the value of the option. On the expiry date the price of the option will be a function of the strike price and the price of the underlying asset. However for pricing purposes an option trader must take into account all the factors above. From Chapter 16 we know that the value of an option is composed of intrinsic value and time value: intrinsic value is apparent straight away when an option is struck, and a valuation model is essentially pricing the time value of the option. This is considered next.

Assessing time value

The *time value* of an option reflects the fact that it is highest for at-the-money options, and also higher for an in-the-money option than an out-of-the-money option. This can be demonstrated by considering the hedge process followed by a market maker in options. An out-of-the-money call option, for instance, presents

the lowest probability of exercise for the market maker, therefore she may not even hedge such a position. There is a risk of course that the price of the underlying will rise sufficiently to make the option in-the-money, in which case the market maker would have to purchase the asset in the market, thereby suffering a loss. This must be considered by the market maker, but deeply out-of-the-money options are often not hedged. So the risk to the market maker is lowest for this type of option, which means that the time value is also lowest for such an option.

An in-the-money call option carries a greater probability that it will be exercised. A market maker writing such an option will therefore hedge the position, either with the underlying asset, with futures contracts or via a *risk reversal*. This is a long or short position in a call that is reversed to the same position in a put by selling or buying the position forward (and vice versa). The risk with hedging using the underlying is that its price will fall, causing the option not to be exercised and forcing the market maker to dispose of the underlying at a loss. However this risk is lowest for deeply in-the-money options, and this is reflected in the time value for such options, which diminishes the more in-the-money the option is.

The highest risk lies in writing an at-the-money option. In fact the majority of over-the-counter (OTC) options are struck at-the-money. The risk level reflects the fact that there is greatest uncertainty with this option, because there is an even chance of it being exercised. The decision on whether to hedge is therefore not as straightforward. As an at-the-money option carries the greatest risk for the market maker in terms of hedging it, the time value for it is the highest.

American options

In Chapter 18 we discussed the B–S and other models in terms of European options, and also briefly referred to a model developed for American options on dividend-paying securities. In theory an American option will have greater value than an equivalent European option, because of the early-exercise option. This added feature implies a higher value for the American option. In theory this is correct, but in practice it carries lower weight because American options are rarely exercised ahead of expiry. A holder of an American option must assess if it is ever optimal to exercise it ahead of the expiry date, and usually the answer to this is 'no'. This is because, by exercising an option, the holder realises only the intrinsic value of the option. However if the option is traded in the market, that is, sold, then the full value will be realised, including the time value. Therefore it is rare for an American option to be exercised ahead of the expiry date; rather, it will be sold in the market to realise full value.

As the chief characteristic that differentiates American options from European options is rarely employed, in practical terms they do not have greater value than European ones. Therefore they have similar values to equivalent European options. However an option pricing model, calculating the probability that an option will be exercised, will determine under certain circumstances that the American option has a higher probability of being exercised and assign it a higher price.

Under certain circumstances it is optimal to exercise American options early. The most significant is when an option has *negative time value*. An option can have

negative time value when, for instance, a European option is deeply in-the-money and very near to maturity. The time value will be a small positive; however the potential value in deferring cash flows from the underlying asset may outweigh this, leading to a negative time value. The best example of this is for a deeply in-the-money option on a futures contract. By deferring its exercise, the opportunity to invest the cash proceeds from the profit on the futures contract (remember, futures are cash settled daily via the margin process) is lost and this is potential interest income foregone. In such circumstances, it would be optimal to exercise an option ahead of its maturity date, assuming it is an American one. Therefore when valuing an American option, the probability of it being exercised early is considered and if it is deeply in-the-money this probability will be at its highest.

Measuring option risk: the Greeks

It should have become apparent from a reading of the previous chapters that the price sensitivity of options is different from that of other financial market instruments. This is clear from the variables that are required when pricing an option, which we presented by way of recap at the start of this chapter. The value of an option is sensitive to changes in one or any combination of the five variables that are used in the valuation.[1] This makes risk managing an option book more complex than managing other instruments. For example, the value of a swap is sensitive to one variable only, the swap rate. The relationship between the change in value of the swap and the change in the swap rate is also a linear one. A bond futures contract is priced as a function of the current spot price of the cheapest-to-deliver bond and the current money market repo rate. Options on the other hand react to moves in any of the variables used in pricing them; more importantly the relationship between the value of the option and the change in a key variable is not a linear one. The market uses a measure for each of the variables, and in some cases for a derivative of these variables. These are termed the 'Greeks' as they are called after letters in the ancient Greek alphabet.[2] In this section we review these sensitivity measures and how they are used.

Delta

The *delta* of an option is a measure of how much the value or premium of the option changes with changes in the price of the underlying asset. That is,

$$\delta = \frac{\Delta C}{\Delta S}$$

[1] Of course, the strike price for a plain vanilla option is constant.

[2] All but one; the term for the volatility sensitivity measure, *vega*, is not a Greek letter. In certain cases one will come across the use of the term *kappa* to refer to volatility, and this is a Greek letter. However it is more common for volatility to be referred to by the term vega.

Mathematically the delta of an option is the partial derivative of the option premium with respect to the underlying, given by (19.1):

$$\delta = \frac{\partial C}{\partial S} \tag{19.1}$$

or

$$\delta = \frac{\partial P}{\partial S}$$

In fact the delta of an option is given by the $N(d_1)$ term in the B–S equation. It is closely related to, but not equal to, the probability that an option will be exercised. If an option has a delta of 0.6 or 60%, this means that a £100 increase in the value of the underlying will result in a £60 increase in the value of the option. Delta is probability the most important sensitivity measure for an option, as it measures the sensitivity of the option price to changes in the price of the underlying, and this is very important for option market makers.

It is also the main hedge measure. When an option market maker wishes to hedge a sold option, she may do this by buying a matching option, by buying or selling another instrument with the same but opposite value as the sold option, or by buying or selling the underlying. If the hedge is put on with the underlying, the amount is governed by the delta. So for instance if the delta of an option written on one ordinary share is 0.6 and a trader writes 1000 call options, the hedge would be a long position in 600 of the underlying shares. This means that if the value of the shares rises by £1, the £600 rise in the value of the shares will offset the £600 loss in the option position. This is known as *delta hedging*. As we shall see later on, this is not a static situation, and the fact that delta changes, and is also an approximation, means that hedges must be monitored and adjusted, so-called *dynamic hedging*.

The delta of an option measures the extent to which the option moves with the underlying asset price; at a delta of zero the option does not move with moves in the underlying, while at a delta of 1 it will behave identically to the underlying.

A positive delta position is equivalent to being long the underlying asset, and can be interpreted as a bullish position. A rise in the asset price results in profit, as in theory a market maker could sell the underlying at a higher price, or in fact sell the option. The opposite is true if the price of the underlying falls. With a positive delta, a market maker would be over-hedged if running a delta-neutral position. Table 19.1 shows the effect of changes in the underlying price on the delta position in the option book; to maintain a delta-neutral hedge, the market maker must buy or sell delta units of the underlying asset, although in practice futures contracts may be used.

Gamma

In a similar way that the modified duration measure becomes inaccurate for larger yield changes, due to the nature of its calculation, there is an element of inaccuracy

Table 19.1 Delta-neutral hedging for changes in underlying price

Option	*Rise in underlying asset price*	*Fall in underlying asset price*
Long call	Rise in delta: sell underlying	Fall in delta: buy underlying
Long put	Fall in delta: sell underlying	Rise in delta: buy underlying
Short call	Rise in delta: buy underlying	Fall in delta: sell underlying
Short put	Fall in delta: buy underlying	Rise in delta: sell underlying

with the delta measurement and with delta hedging an option book. This is because the delta itself is not static, and changes with changes in the price of the underlying. A book that is *delta-neutral* at one level may not remain so as the underlying price changes. To monitor this, option market makers calculate *gamma*. The gamma of an option is a measure of how much the delta value changes with changes in the underlying price. It is given by:

$$\Gamma = \frac{\Delta\delta}{\Delta S}$$

Mathematically gamma is the second partial derivative of the option price with respect to the underlying price, that is:

$$\frac{\partial^2 C}{\partial S^2} \text{ or } \frac{\partial^2 P}{\partial S^2}$$

and is given by (18.2):

$$\Gamma = \frac{N(d_1)}{S\sigma\sqrt{T}} \tag{19.2}$$

The delta of an option does not change rapidly when an option is deeply in or out-of-the-money, so in these cases the gamma is not significant. However when an option is close to or at-the-money, the delta can change very suddenly and at that point the gamma is very large. The value of gamma is positive for long call and put options, and negative for short call and put options. An option with high gamma causes the most problems for market makers, as the delta hedge must be adjusted constantly, which will lead to high transaction costs. The higher the gamma, the greater is the risk that the option book is exposed to loss from sudden moves in the market. A negative gamma exposure is the highest risk, and this can be hedged only by putting on long positions in other options. A perfectly hedged book is gamma neutral, which means that the delta of the book does not change.

When gamma is positive, a rise in the price of the underlying asset will result in a higher delta. Adjusting the hedge will require selling the underlying asset or futures contracts. The reverse applies if there is a fall in the price of the underlying. As the hedge adjustment is made in the same direction as that in which the market

is moving, this adjustment is possibly easier to conceptualise for newcomers to a market-making desk. When adjusting a hedge in a rising market, underlying assets or futures are sold, which in itself may generate profit. In a falling market, the delta hedge is insufficient and must be rebalanced through purchase of the underlying.

However with a negative gamma, an increase in the price of the underlying will reduce the value of the delta, so therefore to adjust the delta hedge, the market maker must buy more of the underlying asset or futures equivalents. However, when the underlying asset price falls, the delta will rise, necessitating selling of the underlying asset to rebalance the hedge. In this scenario, irrespective of whether cash or off-balance sheet instruments are being used, the hedge involves selling assets in a falling market, which will generate losses even as the hedge is being put on. Negative gamma is therefore a high-risk exposure in a rising market.

Managing an option book that has negative gamma is more risky if the underlying asset price volatility is high. In a rising market the market maker becomes short and must purchase more of the underlying, which may produce losses. The same applies in a falling market. If the desk is pursuing a delta-neutral strategy, running a positive gamma position should enable generation of profit in volatile market conditions. Under the same scenario, a negative gamma position would be risky and would be excessively costly in terms of dynamically hedging the book.

Gamma is the only one of the major Greeks that does not measure the sensitivity of the option premium: instead it measures the change in delta. The delta of an option is its hedge ratio, and gamma is a measure of how much this hedge ratio changes for changes in the price of the underlying. This is why a gamma value results in problems in hedging an option book, as the hedge ratio is always changing. This ties in with our earlier comment that at-the-money options have the highest value, because they present the greatest uncertainty and hence the highest risk. The relationship is illustrated by the behaviour of gamma, which follows that of the delta.

To adjust an option book so that it is gamma-neutral, a market maker must put on positions in an option on the underlying or on the future. This is because the gamma of the underlying and the future is zero. It is common for market makers to use exchange-traded options. Therefore a book that needs to be made gamma-neutral must be rebalanced with options; however, by adding to its option position, the book's delta will alter. Therefore to maintain the book as delta-neutral, the market maker will have to rebalance it using more of the underlying asset or futures contracts. The calculation made to adjust gamma is a snapshot in time, and as the gamma value changes dynamically with the market, the gamma hedge must be continually rebalanced, like the delta hedge, if the market maker wishes to maintain the book as gamma-neutral.

Theta

The *theta* of an option measures the extent of the change in value of an option with change in the time to maturity. That is, it is:

$$\Theta = \frac{\Delta C}{\Delta T}$$

or

$$-\frac{\partial C}{\partial T} \text{or} -\frac{\partial P}{\partial T}$$

and, from the formula for the B–S model, mathematically it is given for a call option as (19.3):

$$\Theta = -\frac{S\sigma}{2\sqrt{2\pi T}} e^{\frac{-d_1^2}{2}} - Xre^{-rT}N(d_2) \tag{19.3}$$

Theta is a measure of time decay for an option. A holder of a long option position suffers from time decay because as the option approaches maturity, its value is made up increasingly of intrinsic value only, which may be zero as the option approaches expiry. For the writer of an option, the risk exposure is reduced as a result of time decay, so it is favourable for the writer if the theta is high. There is also a relationship between theta and gamma, however: when an option gamma is high, its theta is also high, and this results in the option losing value more rapidly as it approaches maturity. Therefore a high theta option, while welcome to the writer, has a down-side because it is also high gamma. There is therefore in practice no gain for the writer of an option to be high theta.

The theta value impacts certain option strategies. For example, it is possible to write a short-dated option and simultaneously purchase a longer-dated option with the same strike price. This is a play on the option theta: if the trader believes that the time value of the longer-dated option will decay at a slower rate than the short-dated option, the trade will generate a profit.

Vega

The *vega* of an option measures how much its value changes with changes in the volatility of the underlying asset. It is also known as *epsilon* (ϵ), *eta* (η), or *kappa* (κ).

We define vega as:

$$v = \frac{\Delta C}{\Delta \sigma}$$

or

$$v = \frac{\partial C}{\partial \sigma} \text{ or } \frac{\partial P}{\partial \sigma}$$

and mathematically from the B–S formula it is defined in (19.4) for a call or put.

$$v = \frac{S\sqrt{\frac{T}{2\pi}}}{e^{\frac{d_1^2}{2}}} \tag{19.4}$$

It may also be given by (19.5):

$$v = S\sqrt{\Delta T}N(d_1) \tag{19.5}$$

An option exhibits its highest vega when it is at-the-money, and decreases as the underlying and strike prices diverge. Options with only a short time to expiry have a lower vega compared with longer-dated options. An option with positive vega generally has positive gamma. Vega is also positive for a position composed of long call and put options, and an increase in volatility will then increase the value of the options. A vega of 12.75 means that for a 1% increase in volatility, the price of the option will increase by 0.1275.

Buying options is the equivalent of buying volatility, while selling options is equivalent to selling volatility. Market makers generally like volatility and set up their books so that they are positive vega. The basic approach for volatility trades is that the market maker will calculate the implied volatility inherent in an option price, and then assess whether this is accurate compared with her own estimation of volatility. Just as positive vega is long call and puts, if the trader feels the implied volatility in the options is too high, she will put on a short vega position of short calls and puts, and then reverse the position out when the volatility declines.

Table 19.2 shows the response to a delta hedge following a change in volatility.

Managing an option book involves trade-offs between the gamma and the vega, much like there are between gamma and theta. A long in options means long vega and long gamma, which is not conceptually difficult to manage. However if there is a fall in volatility levels, the market maker can either maintain positive gamma,

Table 19.2 Dynamic hedging as a result of changes in volatility

Option position	*Rise in volatility*	*Fall in volatility*
Long call		
ATM	No adjustment to delta	No adjustment to delta
ITM	Rise in delta, buy underlying	Rise in delta, sell underlying
OTM	Fall in delta, sell underlying	Fall in delta, buy underlying
Long put		
ATM	No adjustment to delta	No adjustment to delta
ITM	Fall in delta, sell underlying	Rise in delta, buy underlying
OTM	Rise in delta, buy underlying	Fall in delta, sell underlying
Short call		
ATM	No adjustment to delta	No adjustment to delta
ITM	Fall in delta, sell underlying	Rise in delta, buy underlying
OTM	Rise in delta, buy underlying	Fall in delta, sell underlying
Short put		
ATM	No adjustment to delta	No adjustment to delta
ITM	Rise in delta, buy underlying	Rise in delta, sell underlying
OTM	Fall in delta, sell underlying	Fall in delta, buy underlying

depending on her view of whether the fall in volatility can be offset by adjusting the gamma in the direction of the market, or she can sell volatility (that is, write options) and set up a position with negative gamma. In either case the costs associated with rebalancing the delta must compensate for the reduction in volatility.

Rho

The *rho* of an option is a measure of how much its value changes with changes in interest rates. Mathematically this is:

$$\frac{\partial C}{\partial r} \text{ or } \frac{\partial P}{\partial r}$$

and the formal definition, based on the B–S model formula, is given as (19.6) for a call option.

$$\rho = Xte^{-rT}N(d_2) \tag{19.6}$$

The level of rho tends to be higher for longer-dated options. It is probably the least used of the sensitivity measures because market interest rates are probably the least variable of all the parameters used in option pricing.

Lambda

The *lambda* of an option is similar to its delta in that it measures the change in option value for a change in underlying price. However lambda measures this sensitivity as a percentage change in the price for a percentage change in the price of the underlying. Hence lambda measures the gearing or leverage of an option. This in turn gives an indication of expected profit or loss for changes in the price of the underlying. From Figure 19.1 we note that in-the-money options have a gearing of a minimum of five, and sometimes the level is considerably higher. This means that if the underlying was to rise in price, the holder of the long call could benefit by a minimum of five times more than if he had invested the same cash amount in the underlying instead of in the option.

This has been a brief review of the sensitivity measures used in managing option books. They are very useful to market makers and portfolio managers because they enable them to see what the impact of changes in market rates is on an entire book. A market maker need take only the weighted sum of the delta, gamma, vega and theta of all the options on the book to see the impact of changes on the portfolio. Therefore the combined effect of changes can be calculated, without having to reprice all the options on the book. The Greeks are also important to risk managers and those implementing value-at-risk systems.

The option smile

Our discussion on the behaviour and sensitivity of options prices will conclude with an introduction to the *option smile*. Market makers calculate a measure known

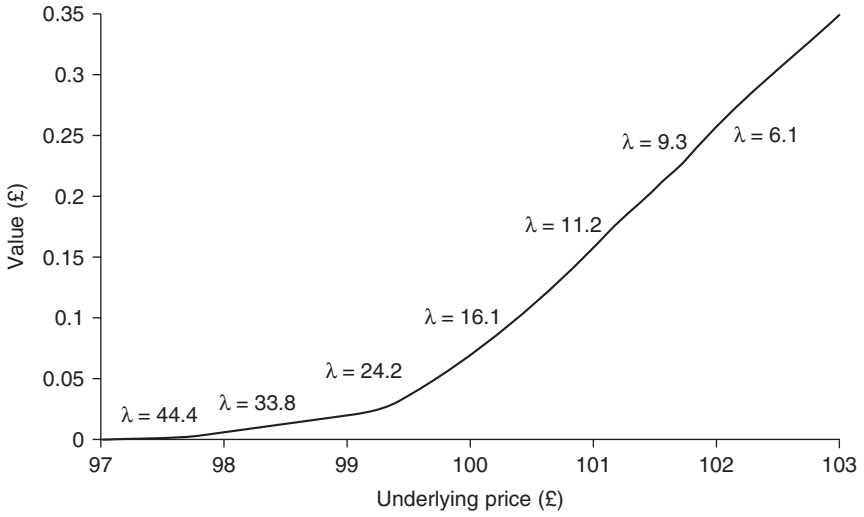


Figure 19.1 Option lambda, nine-month bond option

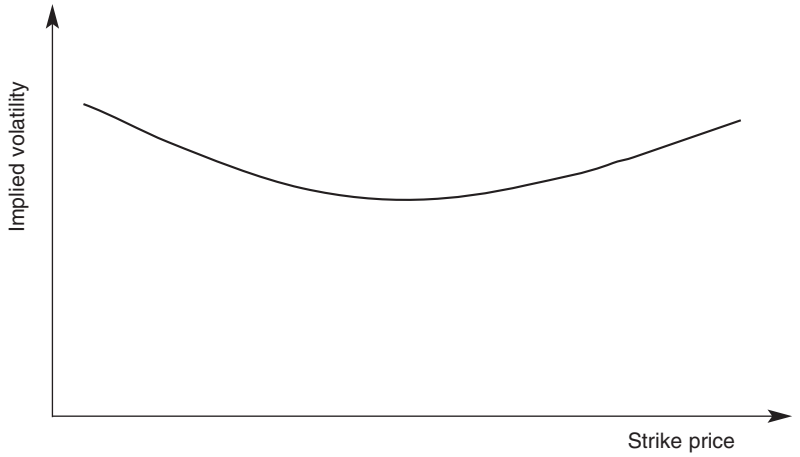


Figure 19.2(a) Bond option volatility smile

as the *volatility smile*, which is a graph that plots the implied volatility of an option as a function of its strike price. The general shape of the smile curve is given in Figure 19.2. What the smile tells us is that out-of-the-money and in-the-money options both tend to have higher implied volatilities than at-the-money options. We define an at-the-money option as one whose strike price is equal to the forward price of the underlying asset.

Under the B–S model assumptions, the implied volatility should be the same across all strike prices of options on the same underlying asset and with the

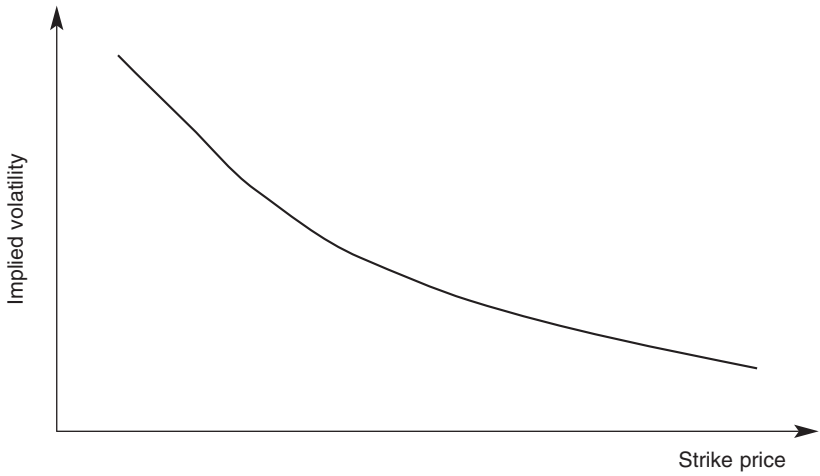


Figure 19.2(b) Equity option volatility smile

same expiry date. However implied volatility is usually observed in the market as a convex function of exercise price, shown in generalised form as Figure 19.2. (In practice, it is not a smooth line or even often a real smile.) The observations confirm that market makers price options with strikes that are less than *S*, and those with strikes higher than *S*, with higher volatilities than options with strikes equal to *S*.

The existence of the volatility smile curve indicates that market makers make more complex assumptions about the behaviour of asset prices than can be fully explained by the geometric Brownian motion model. As a result, market makers attach different probabilities to terminal values of the underlying asset price than those that are consistent with a lognormal distribution. The extent of the convexity of the smile curve indicates the degree to which the market price process differs from the lognormal function contained in the B–S model. In particular the more convex the smile curve, the greater the probability the market attaches to extreme outcomes for the price of the asset on expiry, S_T. This is consistent with the observation that in reality, asset price returns follow a distribution with 'fatter tails' than that described by the lognormal distribution. In addition the direction in which the smile curve slopes reflects the skew of the price process function; a positively sloped implied volatility smile curve results in a price returns function that is more positively skewed than the lognormal distribution. The opposite applies for a negatively sloping curve.

The existence of the smile suggests asset price behaviour that is more accurately described by non-standard price processes, such as the jump diffusion model, or a stochastic volatility, as opposed to constant volatility model.

Considerable research has gone into investigating the smile. The book references in this and the previous chapter are good starting points on this subject.

Caps and floors

Caps and *floors* are options on interest rates. They are commonly written on Libor or another interest rate such as Euribor, the US Prime rate or a commercial paper rate. In this section we review caps, which are essentially calls on an interest rate, while a floor is a put on an interest rate.

A cap is an option contract in which an upper limit is placed on the interest rate payable by the borrower on a cash loan. The limit is known as the *cap level*. The seller of the cap, which is the market-making bank, agrees to pay to the buyer the difference between the cap rate and the higher rate should interest rates rise above the cap level. The buyer of the cap is long the option, and will have paid the cap premium to the seller. Hence a cap is a call option on interest rates. The cash loan may have been taken out before the cap, or indeed with another counterparty, or the cap may have been set up alongside the loan as a form of interest-rate risk management. If a cap is set up in conjunction with a cash loan, the notional amount of the cap will be equal to the amount of the loan. Caps can be fairly long-dated options, for example ten-year caps are not uncommon.

In a typical cap, the cap rate is measured alongside the indexed interest rate at the specified fixing dates. So during its life a cap may be fixed semi-annually with the six-month Libor rate. At the fixing date, if the index interest rate is below the cap level, no payment changes hands and if there is a cash loan involved, the borrower will pay the market interest rate on the loan. If the index rate is fixed above the cap level, the cap seller will pay the difference between the index interest rate and the cap level, calculated for the period of the fix (quarterly, semi-annually, and so on) on the notional amount of the cap. Individual contracts, that is each fixing, during the life of the cap are known as *caplets*. The interest payment on each caplet is given by (19.7):

$$Int = \frac{\max[r - rX, 0] \times (N / B) \times M}{1 + r(N / B)} \tag{19.7}$$

where

r is the interest rate fixing for the specified index
rX is the cap level
M is the notional amount of the cap
B is the day base (360 or 365)
N is the number of days in the interest period (days to the next rate fix).

Similarly to FRAs, any payment made is an upfront payment for the period covered, and so is discounted at the index rate level.

As it is a call option on a specified interest rate, the premium charged by a cap market maker will be a function of the probability that the cap is exercised, based on the volatility of the forward interest rate. Caps are frequently priced using the

Black 76 model. The strike rate is the cap level, while the forward rate is used as the 'price' of the underlying. Using Black's model, the option premium is given by

$$C = \frac{\phi M}{1+\phi rf} \times e^{-rT} \times [rfN(d_1) - rXN(d_2)] \tag{19.8}$$

where

$$d_1 = \frac{\ln(rf / rX)}{\sigma f \sqrt{T}} + \frac{\sigma_f}{2}\sqrt{T}$$
$$d_2 = d_1 - \sigma_f \sqrt{T}$$

and

rf is the forward rate for the relevant term (three-month, six-month, etc)
ϕ is the rate fixing frequency, such as semi-annually or quarterly
σf is the forward rate volatility
T is the time period from the start of the cap to the caplet payment date.

Each caplet can be priced individually, and the total premium payable on the cap is the sum of the caplet prices. The Black model assumes constant volatility, and so banks use later models to price products when this assumption is considered to be materially unrealistic.

A vanilla cap pricing calculator is part of the RATE application software, included in this book.

In the same way as caps and caplets, a *floorlet* is essentially a put option on an interest rate, with a sequence of floorlets being known as a floor. This might be used, for example, by a lender of funds to limit the loss of income should interest rate levels fall. If a firm buys a call and sells a floor, this is known as buying a *collar* because the interest rate payable is bound on the upside at the cap level and on the downside at the floor level. It is possible to purchase a *zero-cost collar* where the premium of the cap and floor elements are identical; this form of interest-rate risk management is very popular with corporates.

Selected references and bibliography

Hull, J. *Options, Futures and Other Derivatives*, 4th edn, Prentice Hall, 2000.
Kolb, R. *Futures, Options and Swaps*, 3rd edn, Blackwell, 2000.
Marshall, J. and Bansal, V. *Financial Engineering*, New York Institute of Finance, 1992.
Tucker, A. *Financial Futures, Options and Swaps*, West Publishing, 1991.

20
Credit Derivatives

This chapter describes *credit derivatives*, instruments that are used to manage credit risk in banking and portfolio management. Credit derivatives exist in a number of forms. We classify these into two main forms, *funded* and *unfunded credit derivatives*, and give a description of each form. We then discuss the main uses of these instruments by banks and portfolio managers. We also consider the main credit events that act as triggering events under which payouts are made on credit derivative contracts.

Introduction

Credit derivatives are financial contracts designed to hedge credit risk exposure by providing insurance against losses suffered because of *credit events*. Credit derivatives also allow investors to gain credit risk exposure to assets without having to actually invest cash.

The principle behind credit derivatives is straightforward. Investors desire exposure to debt that has a risk of defaulting because of the higher returns this offers. However, such exposure brings with it concomitant credit risk. This can be managed with credit derivatives. Alternatively, the exposure itself can be taken on synthetically if, for instance, there are compelling reasons that a cash market position cannot be established. The flexibility of credit derivatives provides users with a number of advantages, and as they are over-the-counter (OTC) products they can be designed to meet specific user requirements.

What constitutes a credit event is defined specifically in the legal documents that describe the credit derivative contract. A number of events may be defined as credit events that fall short of full bankruptcy, administration or liquidation of a company. For instance, credit derivatives contracts may be required to pay out under technical as well as actual default.

A *technical default* is a delay in timely payment of an obligation, or a non-payment altogether. If an obligor misses a payment, by even one day, it is said to be in technical default. This delay may be for operational reasons (and so not really a great worry) or it may reflect a short-term cash flow crisis, such as the Argentina

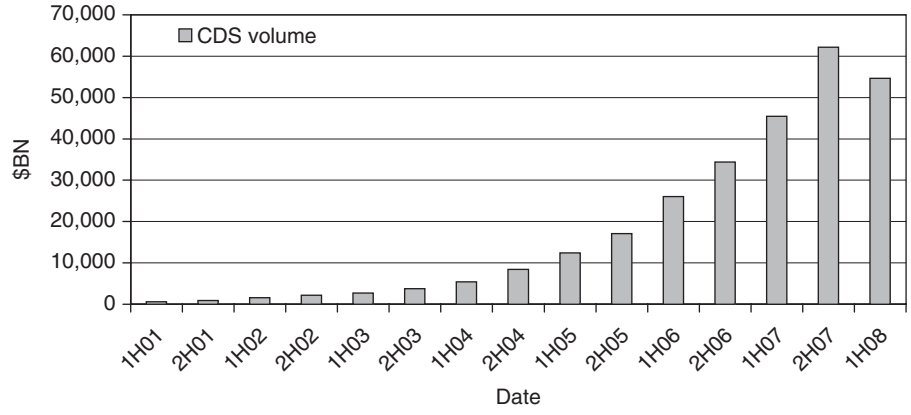


Figure 20.1 Credit default swap notional volume outstanding, 2001–2008

debt default for three months. But if the obligor states it intends to pay the obligation as soon as it can, and specifies a time-span that is within (say) one to three months, then while it is in technical default it is not in actual default. If an obligor is in *actual default*, it is in default and declared as being in default. This does not mean a mere delay of payment. If an obligor does not pay, and does not declare an intention to pay an obligation, it may then be classified by the ratings agencies as being in 'default' and rated 'D'.

If there is a technical or actual default by the borrower so that, for instance, a bond is marked down in price, the losses suffered by the investor can be recouped in part or in full through the payout made by the credit derivative. A payout under a credit derivative is triggered by a credit event. As banks define default in different ways, the terms under which a credit derivative is executed usually include a specification of what constitutes a credit event.

Figure 20.1 illustrates growth in the market, with a contraction following the 2007 credit crunch.

Why use credit derivatives?

Credit derivative instruments enable participants in the financial market to trade in credit as an asset, as they isolate and transfer credit risk. They also enable the market to separate funding considerations from credit risk.

Credit derivatives have two main types of application:

- *Diversifying the credit portfolio*. A bank or portfolio manager may wish to take on credit exposure by providing credit protection in return for a fee. This enhances income on the portfolio. It may sell credit derivatives to enable non-financial counterparties to gain credit exposures, if these clients are unable or unwilling to purchase the assets directly. In this respect the bank or asset manager performs a *credit intermediation* role.

- *Reducing credit exposure*. A bank can reduce credit exposure for either an individual loan or a sectoral concentration by buying a *credit default swap*. This may be desirable for assets that cannot be sold for client relationship or tax reasons. For fixed-income managers, a particular asset or collection of assets may be viewed as an attractive holding in the long term, but at risk from short-term downward price movement. In this instance a sale would not fit in with long-term objectives; however, short-term credit protection can be obtained via a credit swap. For instance, a bank can buy credit protection on a BB-rated entity from a AA-rated bank. It then has eliminated its credit risk to the BB entity, and substituted it for AA-rated counterparty risk. Notice that as the bank retains a counterparty risk to the credit default swap issuer, its credit risk exposure is never completely removed. In practice this is only a serious problem if the bank cannot manage counterparty risk through careful selection and diversification of counterparties.

The intense competition amongst commercial banks, combined with rapid disintermediation, has meant that banks have been forced to evaluate their lending policy with a view to improving profitability and return on capital. The use of credit derivatives assists banks with restructuring their businesses, because they allow banks to repackage and parcel out credit risk, while retaining assets on their balance sheet (when required) and thus maintaining client relationships. As the instruments isolate certain aspects of credit risk from the underlying loan or bond and transfer them to another entity, it becomes possible to separate the ownership and management of credit risk from the other features of ownership of the assets in question. This means that illiquid assets such as bank loans and illiquid bonds can have their credit risk exposures transferred; the bank owning the assets can protect itself against credit loss even if it cannot transfer the assets themselves.

The same principles carry over to the credit risk exposures of portfolio managers. For fixed-income portfolio managers, some of the advantages of credit derivatives are:

- They can be tailor-made to meet the specific requirements of the entity buying the risk protection, as opposed to the liquidity or term of the underlying reference asset.
- They can be 'sold short' without risk of a liquidity or delivery squeeze, as it is a specific credit risk that is being traded. In the cash market it is not possible to 'sell short' a bank loan, for example, but a credit derivative can be used to establish synthetically the same economic effect.
- As they theoretically isolate credit risk from other factors such as client relationships and interest-rate risk, credit derivatives introduce a formal pricing mechanism to price credit issues only. This means a market can develop in credit only, allowing more efficient pricing; it even becomes possible to model a term structure of credit rates.
- When credit derivatives are embedded in certain fixed-income products, such as structured notes and credit-linked notes, they are then off-balance sheet instruments (albeit part of a structure that may have on-balance sheet elements) and

as such incorporate tremendous flexibility and leverage, exactly like other financial derivatives. For instance, bank loans are not particularly attractive investments for certain investors because of the administration required in managing and servicing a loan portfolio. However, an exposure to bank loans and their associated return can be achieved by a total return swap, for instance, while simultaneously avoiding the administrative costs of actually owning the assets. Hence, credit derivatives allow investors access to specific credits while allowing banks access to further distribution for bank loan credit risk.

- They enable institutions to take a view on credit positions to take advantage of perceived anomalies in the price of secondary market loans and bonds, and the price of credit risk.

Thus credit derivatives can be an important instrument for bond portfolio managers as well as commercial banks wishing to increase the liquidity of their portfolios, gain from the relative value arising from credit pricing anomalies, and enhance portfolio returns.

Classification of credit derivative instruments

A number of instruments come under the category of credit derivatives. Irrespective of the particular instrument under consideration, all credit derivatives can be described using the following characteristics:

- the *reference entity*, which is the asset or name on which credit protection is being bought and sold
- the *credit event*, or events, which indicate that the reference entity is experiencing or about to experience financial difficulty and which act as trigger events for payments under the credit derivative contract
- the *settlement mechanism* for the contract, whether cash settled or physically settled
- the *deliverable obligation* that the protection buyer delivers (under physical settlement) to the protection seller on the occurrence of a trigger event.

As we noted earlier, credit derivatives are grouped into *funded* and *unfunded* instruments. In a funded credit derivative, typified by a *credit-linked note* (CLN), the investor in the note is the credit-protection seller and is making an upfront payment to the protection buyer when it buys the note. Thus, the protection buyer is the issuer of the note. If no credit event occurs during the life of the note, the redemption value of the note is paid to the investor on maturity. If a credit event does occur, then on maturity a value less than par will be paid out to the investor. This value will be reduced by the nominal value of the reference asset to which the CLN is linked. The exact process will differ according to whether *cash settlement* or *physical settlement* has been specified for the note. We consider this later.

In an unfunded credit derivative, typified by a *credit default swap* (CDS), the protection seller does not make an upfront payment to the protection buyer. Instead, the protection seller will pay the nominal value of the contract (the amount

insured, in effect), on occurrence of a credit event, minus the current market value of the asset or its recovery value.

Definition of a credit event

The occurrence of a specified credit event will trigger the default payment by the protection seller to the protection buyer. Contracts specify physical or cash settlement. In physical settlement, the protection buyer transfers to the protection seller the deliverable obligation (usually the reference asset or assets), with the total principal outstanding equal to the nominal value specified in the default swap contract. The protection seller simultaneously pays to the buyer 100% of the nominal value. In cash settlement, the protection seller hands to the buyer the difference between the nominal amount of the default swap and the final value for the same nominal amount of the reference asset. This final value is usually determined by means of a poll of dealer banks.

The following may be specified as credit events in the legal documentation between counterparties:

- downgrade in S&P and/or Moodys credit rating below a specified minimum level
- financial or debt restructuring, for example occasioned under administration or as required under US bankruptcy protection
- bankruptcy or insolvency of the reference asset obligor
- default on payment obligations such as bond coupon and continued non-payment after a specified time period
- technical default, for example the non-payment of interest or coupon when it falls due
- a change in credit spread payable by the obligor above a specified maximum level.

The 1999 International Swaps and Derivatives Association (ISDA) credit default swap documentation specifies bankruptcy, failure to pay, obligation default, debt moratorium and 'restructuring' to be credit events. Note that it does not specify a rating downgrade to be a credit event.[1]

The precise definition of 'restructuring' is open to debate, and has resulted in legal disputes between protection buyers and sellers. Prior to issuing its 1999 definitions, ISDA had specified restructuring as an event or events that resulted in making the terms of the reference obligation 'materially less favourable' to the creditor (or protection seller) from an economic perspective. This definition is open to more than one interpretation, and caused controversy when determining if a credit event had occurred. The 2001 definitions specified more precise conditions, including any action that resulted in a reduction in the amount of principal. In the European

[1]The ISDA definitions from 1999 and restructuring supplement from 2001 are available at www.ISDA.org

market, restructuring is generally retained as a credit event in contract documentation, but in the US market it is less common to see it included. Instead, US contract documentation tends to include as a credit event a form of modified restructuring, the impact of which is to limit the options available to the protection buyer as to the type of assets it could deliver in a physically settled contract.

Credit default swaps

The most common credit derivative is the CDS. This is sometimes described as a *credit swap* or *default swap*. A CDS is a bilateral contract in which a periodic fixed fee or a one-off premium is paid to a protection seller, in return for which the seller will make a payment on the occurrence of a specified credit event. The fee is usually quoted as a basis point multiplier of the nominal value. It is usually paid quarterly in arrears.

The swap can refer to a single asset, known as the *reference asset* or *underlying asset*, or a basket of assets. The default payment can be paid in whatever way suits the protection buyer or both counterparties. For example, it may be linked to the change in price of the reference asset or another specified asset, it may be fixed at a predetermined recovery rate, or it may be in the form of actual delivery of the reference asset at a specified price. The basic structure is illustrated in Figure 20.2.

The maturity of the credit swap does not have to match the maturity of the reference asset, and often does not. On occurrence of a credit event, the swap contract is terminated and a settlement payment made by the protection seller or guarantor to the protection buyer. This termination value is calculated at the time of the credit event, and the exact procedure that is followed to calculate the termination value will depend on the settlement terms specified in the contract. This will be either cash settlement or physical settlement:

- *Cash settlement*. The contract may specify a predetermined payout value on occurrence of a credit event. This may be the nominal value of the swap contract. Such a swap is known in some markets as a *digital credit derivative*. Alternatively, the termination payment is calculated as the difference between the

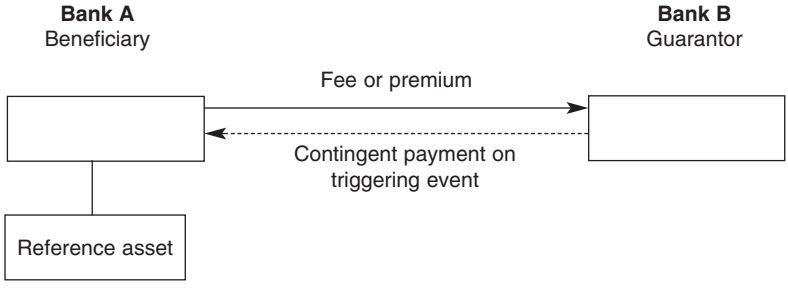


Figure 20.2 Credit default swap

nominal value of the reference asset and its market value at the time of the credit event. This arrangement is more common with cash-settled contracts.[2]

- *Physical settlement.* On occurrence of a credit event, the buyer delivers the reference asset to the seller, in return for which the seller pays the face value of the delivered asset to the buyer. The contract may specify a number of alternative assets that the buyer can deliver; these are known as *deliverable obligations*. This may apply when a swap has been entered into on a reference name rather than a specific obligation (such as a particular bond) issued by that name. Where more than one deliverable obligation is specified, the protection buyer will invariably deliver the asset that is the cheapest on the list of eligible assets. This gives rise to the concept of the *cheapest to deliver*, as encountered with government bond futures contracts, and is in effect an embedded option afforded to the protection buyer.

In theory, the value of protection is identical irrespective of which settlement option is selected. However, under physical settlement the protection seller can gain if there is a recovery value that can be extracted from the defaulted asset; or its value may rise as the fortunes of the issuer improve. Despite this, swap market-making banks often prefer cash settlement as there is less administration associated with it. It is also more suitable when the swap is used as part of a synthetic structured product, because such vehicles may not be set up to take delivery of physical assets.

Another advantage of cash settlement is that it does not expose the protection buyer to any risks should there not be any deliverable assets in the market, for instance due to shortage of liquidity in the market. Were this to happen, the buyer might find the value of its settlement payment reduced. Nevertheless, physical settlement is widely used because counterparties wish to avoid the difficulties associated with determining the market value of the reference asset under cash settlement. Physical settlement also permits the protection seller to take part in the creditor negotiations with the reference entity's administrators, which may result in improved terms for them as holders of the asset.

Example 20.1

XYZ plc credit spreads are currently trading at 120 basis points (bps) relative to government-issued securities for five-year maturities and 195 bps for ten-year maturities. A portfolio manager hedges a $10 million holding of 10-year paper by purchasing the following CDS, written on the five-year bond. This hedge protects for the first five years of the holding, and in the event of XYZ's credit spread widening, will increase in value and may be

[2] Determining the market value of the reference asset at the time of the credit event may be a little problematic as the issuer of the asset may well be in default or administration. An independent third-party calculation agent is usually employed to make the termination payment calculation.

sold on before expiry at profit. The ten-year bond holding also earns 75 bps over the shorter-term paper for the portfolio manager.

Term	*5 years*
Reference credit	XYZ plc five-year bond
Credit event payout date	The business day following occurrence of specified credit event
Default payment	Nominal value of bond × (100 − price of bond after credit event)
Swap premium	3.35%.

Assume now that midway into the life of the swap there is a technical default on the XYZ plc five-year bond, such that its price now stands at \$28. Under the terms of the swap the protection buyer delivers the bond to the seller, who pays out \$7.2 million to the buyer.

The CDS enables one party to transfer its credit risk exposure to another party. Banks may use default swaps to trade sovereign and corporate credit spreads without trading the actual assets themselves; for example, someone who has gone long a default swap (the protection buyer) will gain if the reference asset obligor suffers a rating downgrade or defaults, and can sell the default swap at a profit if he can find a buyer counterparty.[3] This is because the cost of protection on the reference asset will have increased as a result of the credit event. The original buyer of the default swap need never have owned a bond issued by the reference asset obligor.

Credit default swaps are used extensively for flow trading (that is, the daily customer buy and sell business) of single reference name credit risks or, in *portfolio swap* form, for trading a basket of reference credits. CDSs and CLNs are also used in structured products, in various combinations, and their flexibility has been behind the growth and wide application of the synthetic collateralised debt obligation and other credit hybrid products.

Figure 20.3 shows sovereign CDS prices during 2008–9. The graph shows the level of fluctuation in CDS prices.

[3] Be careful with terminology here. To 'go long' of an instrument generally is to purchase it. In the cash market, going long of the bond means one is buying the bond and so receiving coupon; the buyer has therefore taken on credit risk exposure to the issuer. In a CDS, to go long is to buy the swap, but the buyer is purchasing protection and therefore paying premium; the buyer has no credit exposure on the name and has in effect 'gone short' on the reference name (the equivalent of shorting a bond in the cash market and paying coupon). So buying a CDS is frequently referred to in the market as 'shorting' the reference entity.

Figure 20.3 Sovereign CDS levels during 2008–9
© Bloomberg L.P. Used with permission. Visit www.bloomberg.com

Credit-linked notes

A standard CLN is a security, usually issued by an investment-grade entity, that has an interest payment and fixed maturity structure similar to a vanilla bond. The performance of the note, however, including the maturity value, is linked to the performance of a specified underlying asset or assets, as well as to that of the issuing entity. Notes are usually issued at par. The notes are often used by borrowers to hedge against credit risk, and by investors to enhance the yield received on their holdings. Hence, the issuer of the note is the protection buyer and the buyer of the note is the protection seller.

CLNs are essentially hybrid instruments that combine a credit derivative with a vanilla bond. The CLN pays regular coupons; however, the credit derivative element is usually set to allow the issuer to decrease the principal amount if a credit event occurs. For example, consider an issuer of credit cards that wants to fund its (credit card) loan portfolio via an issue of debt. In order to hedge the credit risk of the portfolio, it issues a two-year CLN. The principal amount of the bond is 100% as usual, and it pays a coupon of 7.50%, which is 200 bps above the two-year benchmark. If, however, the incidence of bad debt amongst credit card holders exceeds 10%, the terms state that note holders will only receive back £85 per £100 nominal. The credit-card issuer has in effect purchased a credit option that lowers its liability in the event that it suffers from a specified credit event, which in this

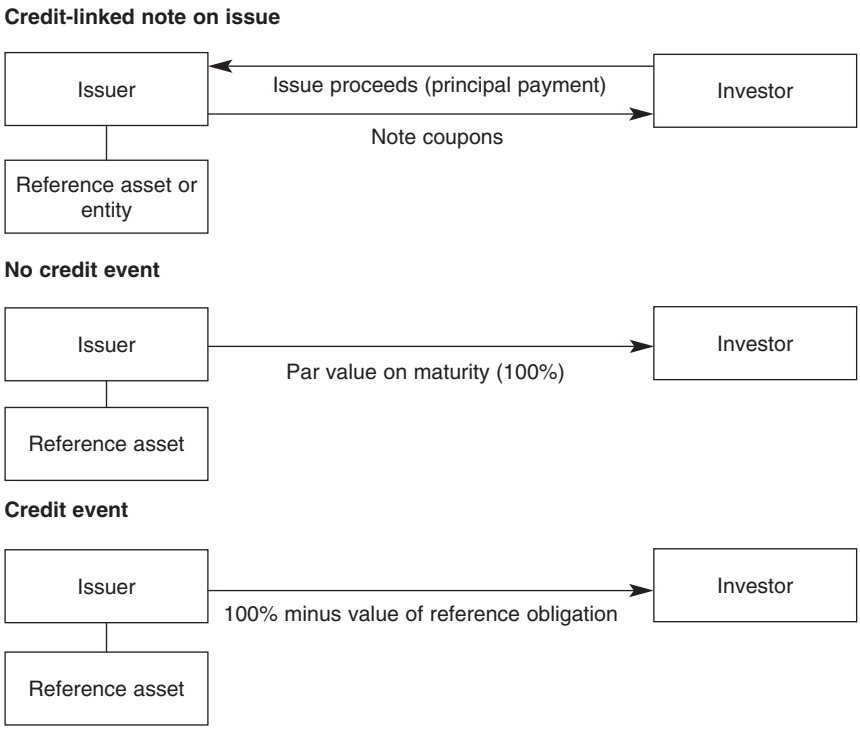


Figure 20.4 A cash-settled credit-linked note

case is an above-expected incidence of bad debts. The credit-card bank has issued the CLN to reduce its credit exposure, in the form of this particular type of credit insurance. If the incidence of bad debts is low, the note is redeemed at par. However, if there a high incidence of such debt, the bank will only have to repay a part of its loan liability.

Figure 20.4 depicts the cash flows associated with a CLN. CLNs exist in a number of forms, but all of them contain a link between the return they pay and the credit-related performance of the underlying asset. Investors may wish to purchase the CLN because the coupon paid on it will be above what the same bank would pay on a vanilla bond it issued, and higher than other comparable investments in the market. In addition, such notes are usually priced below par on issue. Assuming the notes are eventually redeemed at par, investors will also have realised a substantial capital gain.

Total return swaps

A *total return swap* (TRS), sometimes known as a *total rate of return swap* or *TR swap,* is an agreement between two parties to exchange the total returns from financial assets. This is designed to transfer the credit risk from one party to the other. It is

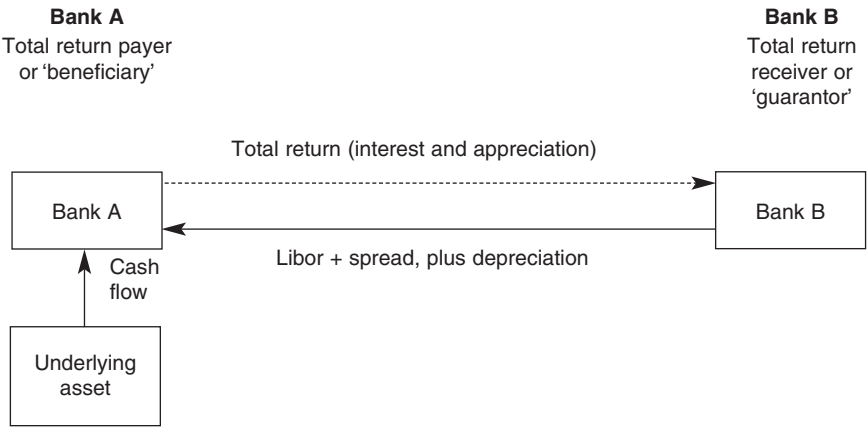


Figure 20.5 Total return swap

one of the principal instruments used by banks and other financial instruments to manage their credit risk exposure, and as such is a credit derivative. One definition of a TRS is given in Francis, Frost and Whittaker (1999), which states that a TRS is a swap agreement in which the *total return* of a bank loan or credit-sensitive security is exchanged for some other cash flow, usually tied to Libor or some other loan or credit-sensitive security.

In some versions of a TRS the actual underlying asset is sold to the counterparty, with a corresponding swap transaction agreed alongside; in other versions there is no physical change of ownership of the underlying asset. The TRS trade itself can be to any maturity term, that is, it need not match the maturity of the underlying security. In a TRS the total return from the underlying asset is paid over to the counterparty in return for a fixed or floating cash flow. This makes it slightly different from other credit derivatives, as the payments between counterparties to a TRS are connected to changes in the market value of the underlying asset, as well as changes resulting from the occurrence of a credit event.

Figure 20.5 illustrates a generic TR swap. The two counterparties are labelled as banks, but the party termed 'Bank A' can be another financial institution, including insurance companies and hedge funds that often hold fixed income portfolios. In Figure 20.5 Bank A has contracted to pay the 'total return' on a specified reference asset, while simultaneously receiving a Libor-based return from Bank B. The reference or underlying asset can be a bank loan such as a corporate loan or a sovereign or corporate bond. The total return payments from Bank A include the interest payments on the underlying loan as well as any appreciation in the market value of the asset. Bank B will pay the Libor-based return; it will also pay any difference if there is a depreciation in the price of the asset. The economic effect is as if Bank B owned the underlying asset, and as such TR swaps are synthetic loans or securities. A significant feature is that Bank A will usually hold the underlying asset on its balance sheet, so that if this asset was originally on Bank B's balance sheet, this

is a means by which the latter can have the asset removed from its balance sheet for the term of the TR swap.[4] If we assume Bank A has access to Libor funding, it will receive a spread on this from Bank B. Under the terms of the swap, Bank B will pay the difference between the initial market value and any depreciation, so it is sometimes termed the 'guarantor', while Bank A is the 'beneficiary'.

The total return on the underlying asset comprises the interest payments and any change in the market value if there is capital appreciation. The value of an appreciation may be settled in cash, or alternatively there may be physical delivery of the reference asset on maturity of the swap, in return for a payment of the initial asset value by the total return 'receiver'. The maturity of the TR swap need not be identical to that of the reference asset, and in fact it is rare for it to be so.

The swap element of the trade will usually pay on a quarterly or semi-annual basis, with the underlying asset being revalued or marked-to-market on the refixing dates. The asset price is usually obtained from an independent third party source such as Bloomberg or Reuters, or as the average of a range of market quotes. If the obligor of the reference asset defaults, the swap may be terminated immediately, with a net present value payment changing hands according to this value, or it may be continued with each party making appreciation or depreciation payments as appropriate. This second option is only available if there is a market for the asset, which is unlikely in the case of a bank loan. If the swap is terminated, each counterparty will be liable to the other for accrued interest plus any appreciation or depreciation of the asset. Commonly under the terms of the trade, the guarantor bank has the option to purchase the underlying asset from the beneficiary bank, then deal directly with the loan defaulter.

With a TRS the basic concept is that one party 'funds' an underlying asset and transfers the total return of the asset to another party, in return for a (usually) floating return that is a spread to Libor. This spread is a function of:

- the credit rating of the swap counterparty
- the amount and value of the reference asset
- the credit quality of the reference asset
- the funding costs of the beneficiary bank
- any required profit margin
- the capital charge associated with the TR swap.

The TRS counterparties must therefore consider a number of risk factors associated with the transaction, which include:

- the probability that the TR beneficiary may default while the reference asset has declined in value
- the reference asset obligor defaults, followed by default of the TR swap receiver before payment of the depreciation has been made to the payer or 'provider'.

[4] Although it is common for the receiver of the Libor-based payments to have the reference asset on its balance sheet, this is not always the case.

The first risk measure is a function of the probability of default by the TRS receiver and the market volatility of the reference asset, while the second risk is related to the joint probability of default of both factors as well as the recovery probability of the asset.

TRS contracts are used in a variety of applications by banks, other financial institutions and corporates. They can written as pure exchanges of cash flow differences – rather like an interest-rate swap – or the reference asset can be actually transferred to the total return payer, which would then make the TRS akin to a 'synthetic repo' contract.[5]

- *As pure exchanges of cash flow differences*. Using TRSs as a credit derivative instrument, a party can remove exposure to an asset without having to sell it. This is conceptually similar to interest-rate swaps, which enable banks and other financial institutions to trade interest-rate risk without borrowing or lending cash funds. A TRS agreement entered into as a credit derivative is a means by which banks can take on unfunded off-balance sheet credit exposure. Higher-rated banks that have access to LIBID funding can benefit by funding on-balance sheet assets that are credit protected through a credit derivative such as a TRS, assuming the net spread of asset income over credit protection premium is positive.
- *Reference asset transferred to the total return payer*. In a vanilla TRS the total return payer retains rights to the reference asset, although in some cases servicing and voting rights may be transferred. The total return receiver gains an exposure to the reference asset without having to pay out the cash proceeds that would be required to purchase it. As the maturity of the swap rarely matches that of the asset, the swap receiver may gain from the positive funding or *carry* that derives from being able to roll over short-term funding of a longer-term asset.[6] The total return payer, on the other hand, benefits from protection against market and credit risk for a specified period of time, without having to liquidate the asset itself. On maturity of the swap the total return payer may reinvest the asset if it continues to own it, or it may sell the asset in the open market. Thus the instrument may be considered a *synthetic repo*.

The economic effect of the two applications may be the same, but they are considered different instruments:

- The TRS as a credit derivative instrument actually takes the assets off the balance sheet, whereas the tax and accounting authorities treat repo as if the assets remain on the balance sheet.
- A TRS trade is conducted under the ISDA standard legal agreement, while repo is conducted under a standard legal agreement called the Global Master Repurchase Agreement (GMRA).

[5] When a bank sells stock short, it must borrow the stock to deliver it to its customer, in return for a fee (called a stock loan), or it may lend cash against the stock which it then delivers to the customer (called a 'sale and repurchase agreement' or *repo*). The counterparty is 'selling and buying back' while the bank that is short the stock is 'buying and selling back'. A TRS is a synthetic form of repo, as the bond is sold to the TRS payer.

[6] This assumes a positively sloping yield curve.

It is these differences that, under certain circumstances, make the TRS funding route a more favourable one.

We now explain in more detail the main uses of TRSs.

Reduction in credit risk

A TRS conducted as a synthetic repo is usually undertaken to effect the temporary removal of assets from the balance sheet. This can be done by entering into a short-term TRS with, say, a two-week term that straddles the reporting date. Bonds are removed from the balance sheet if they are part of a sale plus TRS transaction. This is because legally the bank selling the asset is not required to repurchase bonds from the swap counterparty, nor is the total return payer obliged to sell the bonds back to the counterparty (or indeed sell the bonds at all on maturity of the TRS).

Hence, under a TRS an asset such as a bond position may be removed from the balance sheet. This may be desired for a number of reasons, for example if the institution is due to be analysed by credit rating agencies or if the annual external audit is due shortly. Another reason that a bank may wish to temporarily remove lower-credit-quality assets from its balance sheet is if it is in danger of breaching capital limits between the quarterly return periods. In this case, as the return period approaches, lower-quality assets may be removed from the balance sheet by means of a TRS, which is set to mature after the return period has passed. In summary, to avoid adverse impact on regular internal and external capital and credit exposure reporting, a bank may use TRSs to reduce the amount of lower-quality assets on the balance sheet.

The TRS as a funding instrument

A TRS can be regarded as a funding instrument, in other words as a substitute for a repo trade. There may be legal, administrative, operational or other reasons why a repo trade is not entered into to begin with. In these cases, provided that a counterparty can be found and the funding rate is not prohibitive, a TRS may be just as suitable.

Consider a financial institution such as a regulated broker-dealer that has a portfolio of assets on its balance sheet for which it needs to obtain funding. These assets are investment-grade structured finance bonds such as credit card asset-backed securities, residential mortgage-backed securities and collateralised debt obligation notes, and investment-grade convertible bonds. In the repo market, it is able to fund these at Libor plus 6 bps. That is, it can repo the bonds out to a bank counterparty, and will pay Libor plus 6 bps on the funds it receives.

Assume that for operational reasons the bank can no longer fund these assets using repo. It can fund them using a basket TRS instead, provided a suitable counterparty can be found. Under this contract, the portfolio of assets is swapped out to the TRS counterparty, and cash received from the counterparty. The assets are therefore sold off the balance sheet to the counterparty, an investment bank. The investment bank will need to fund this itself: it might have a line of credit from a parent bank or it might swap the bonds out itself. The funding rate it charges the broker-dealer will depend on the rate at which it can fund the assets itself. Assume

this is Libor plus 12 bps – the higher rate reflects the lower liquidity in the basket TRS market for non-vanilla bonds.

The broker-dealer enters into a three-month TRS with the investment bank counterparty, with a one-week interest-rate reset. This means that at each week interval the basket is revalued. The difference in value from the last valuation is paid (if higher) or received (if lower) by the investment bank to the broker-dealer; in return the broker-dealer also pays one week's interest on the funds it received at the start of the trade. In practice these two cash flows are netted off and only one payment changes hands, just as in an interest-rate swap. The terms of the trade are shown below.

Trade date	22 December 2003
Value date	24 December 2003
Maturity date	24 March 2004
Rate reset	31 December 2003
Interest rate	1.19875% (this is the one-week USD Libor fix of 1.07875 plus 12 bps)

The swap is a three-month TRS with one-week reset, which means that the swap can be broken at one-week intervals and bonds in the reference basket can be returned, added to or substituted.

Assume that the portfolio basket contains five bonds, all US dollar denominated. Assume further that these are all investment-grade credit card asset-backed securities with prices available on Bloomberg. The combined market value of the entire portfolio is taken to be US$151,080,951.00.

At the start of the trade, the five bonds are swapped out to the investment bank, which pays the portfolio value for them. On the first reset date, the portfolio is revalued and the following calculations confirmed:

Old portfolio value	$151,080,951.00
Interest rate	1.19875%
Interest payable by broker-dealer	$35,215.50
New portfolio value	$152,156,228.00
Portfolio performance	+ $1,075,277
Net payment: broker-dealer receives	$1,040,061.50

The rate is reset for value on 31 December 2003 for the period to 7 January 2004. The rate is 12 bps over the one-week USD Libor fix on 29 December 2003, which is 1.15750 + 0.12 or 1.2775%. This interest rate is payable on the new 'loan' amount of $152,156,228.00.

The TRS trade has become a means by which the broker-dealer can obtain collateralised funding for its portfolio. Like a repo, the bonds are taken off the broker-dealer's balance sheet, but unlike a repo the tax and accounting treatment also assumes they have been permanently taken off the balance sheet. In addition, the TRS is traded under the ISDA legal definitions, compared with a repo which is traded under the GMRA standard repo legal agreement.

CDS portfolio indices

Since 2005 CDS Index products have given investors the opportunity to express views on portfolios of credits across world regions and also by type of credit (e.g., investment grade, high yield and cross over). Investors who wish to go long credit risk can sell protection on the relevant CDS Index (the market terminology for this is that the investor is a buyer of the CDS Index).

'On the run' CDS Index contracts are quoted by major credit derivative dealers. The contracts are an OTC bilateral agreement. To enter into the contract an initial upfront payment occurs if there is a difference between the spread of the index and the fixed coupon of the 'on the run' index. For example if the current level of the index spread is lower than the 'on the run' index fixed coupon, the investor who is going long credit risk will make a payment to the seller. The rationale behind this payment is to offset the difference between the fixed coupon and spread and reflects the credit improvement in the position since the 'on the run' index was issued. If the spread was higher than the fixed coupon then the investor would receive the upfront. The level of the index is determined by supply and demand.

Figure 20.6 is an overview of the main index structures.

Overview of CDS Portfolio Indices – by geographical region

North America (NA)
- CDX NA IG 125 Investment Grade credits
- CDX NA XO 35 Crossover credits
- CDX NA HY 100 High Yield credits

Europe
- iTraxx Europe 125 Investment Grade credits
- iTraxx Europe Crossover 40 Crossover credits

Asia (ex-Japan)
- iTraxx Japan 50 Investment Grade credits

Japan
- iTraxx Japan 50 Investment Grade credits

Emerging Markets
- CDX EM 15 Sovereign credits
- CDX EM Diversified 40 Sovereign and corporate credits

Typical Index construction
- Equally weighted
- Most issues 'roll' every 6 months the newly issued index is called 'on the run'.
- 'On the run' indices are more liquid than the older 'off the run' indices
- Issue dates on 20 March, 20 June, 20 September or 20 December.
- The Premium leg pays a quarterly payment of the fixed coupon on the notional on these dates. The accrued premium is calculated on an Act/360 basis
- The most liquid index maturity usually is 5 years, but other maturities can also betraded include 3, 7 and 10 year
- For the CDX index credit events do not include restructuring
- Owner and administrator of the CDX and iTraxx indices is Markit

Figure 20.6 Overview of CDS Portfolio Indices, by geographical region

If a credit event occurs at time T since the last coupon payment date, the investor who is long credit risk pays out an amount equal to $(1 - R)/N$ of the notional value of the index. Where N is the total number of credits in the portfolio and R is the recovery rate of the credit. After a credit event the notional of the index position is written down by $1/N$ amount. As a result the future index coupon payment received by the investor, who is long credit risk, is then based on a reduced notional and is $[(N - 1)/N]^*C$. The long credit risk investor will also receive the fraction of the index coupon which accrued from the previous coupon date. If T is the fraction of the year (actual/360) since the last coupon date and the default date, the accrued coupon is $T^* C/N$, where C is the coupon of the 'on the run' index.

For example, assume the notional is \$1.25 bn and the index is the iTraxx Europe contract. The fixed coupon is 80 bps and the index spread is 90 bps. The seller of protection will receive an upfront premium to reflect the fact that the index spread is higher than the fixed coupon, and the rationale for the upfront is that the credit index has widened. The payment is due to the party who is going long credit risk to reflect the fact that there has been credit deterioration since the date of the issuance of the 'on the run' index.

The amount of this upfront payment to the long credit risk party is calculated as follows:

Notional multiplied by (90–80) bps multiplied by the *RPV*01
(for example, assume this is 4.5)
= \$1.25 bn*10 bps *4.5 = \$5,625,000,

where *RPV*01 is the term used to describe the present value of a \$1 p.a. credit-risky annuity for the term of the remaining term of the CDS. The credit risk is based on the survival curve of the index.

Assuming that a credit event occurs and settles after two months from the date of the last coupon payment, assume the time period T on an actual/360 basis is 61/360, and assume the recovery rate on the credit is 40%.

The seller of protection pays \$1.25 bn/125 * (1–40%) = \$6,000,000 and receives an accrued coupon of \$1.25 bn/125 * 61/360* 80 bps = \$13,556.

The long credit risk party will receive future coupons based on a reduced notional of \$1.25 bn *(124/125) = \$1.24 bn.

So assuming there are no further defaults the next coupon payable after one quarter will be \$1.24 bn* 91/360* 80 bps = \$2,507,556.

Index tranche market

After the introduction of indices on credit portfolios, investors became interested in identifying different risk/reward profiles (as in a synthetic CDO) of the standardised CDS portfolio indices. This investor demand led to the formation of the tranche market. The tranche market is now the most common way for an investor to gain exposure to leveraged corporate credit risk. Figures 20.7 and 20.8 illustrate the product.

The index tranche market trades in both a standardised format and a bespoke format. Bespoke format means that investors can 'customise' their investment by

CDX NA IG

	Attachment point	Detachment point
Super Senior	15%	30%
Senior	10%	15%
Senior Mezzanine	7%	10%
Junior Mezzanine	3%	7%
Equity	0%	3%

CDX NA HY

	Attachment point	Detachment point
Super Senior	35%	60%
Senior	25%	35%
Senior Mezzanine	15%	25%
Junior Mezzanine	10%	15%
Equity	0%	10%

iTraxx Europe

	Attachment point	Detachment point
Super Senior	12%	22%
Senior	9%	12%
Senior Mezzanine	6%	9%
Junior Mezzanine	3%	6%
Equity	0%	3%

Comments:

- Attachment point is the point at which realised credit losses from the collateral pool start to cause realised losses to the tranche
- Detachment point is the point at which realised credit losses from the collateral pool have completely eroded the tranche

Figure 20.7 Overview of Standardised Index Tranche Market

selecting the portfolio constituents (reference credits, number of credits, sector diversification etc), the attachment and detachment points, currency, maturity and target rating.

The index tranche product is an OTC derivative market and a pure synthetic form of corporate credit CDO exposure. The tranche market has been established so that the trade is a bilateral contract between the dealer and the investor. The term 'single-tranche synthetic CDO' is often used to describe an index tranche exposure. However, as a result of the contract the dealer is exposed to the risk (credit, market and operational) of managing the contract.

Correlation trading is the term generally used to describe the activities of banks and hedge funds who participate in the tranche (standardised or bespoke) market. The term is used because the risk management of tranche exposure is highly dependent on the timing and ordering of defaults ('default correlation') of the underlying portfolio credit.

The payoff on the index tranche product is driven by the amount of realised portfolio loss that has eroded the tranche width. The tranche exposure is defined by the attachment point and detachment point. The attachment point is a lower

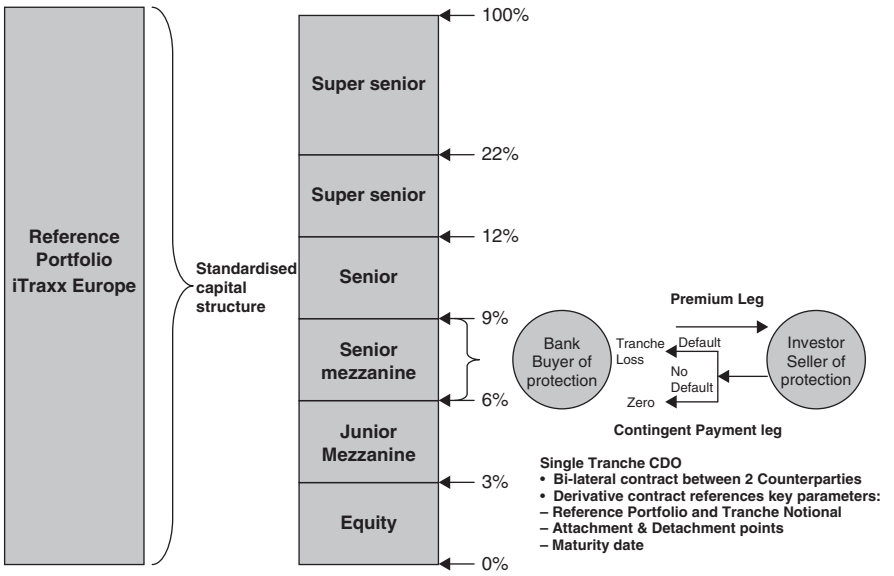


Figure 20.8 Illustration of tranched iTraxx product

percentage than the detachment point and the difference is referred to as the tranche width. For reference portfolio losses that are below the attachment point there is no realised loss to the tranche notional. If losses exceed the detachment point then the tranche notional is completely eroded. When portfolio losses lie between the attachment and detachment point there is a fractional loss to the index tranche. The buyer of protection (short credit risk) pays a premium or coupon leg that is based on the notional outstanding of the tranche. The seller of protection (long credit risk) makes contingent payments dependent on the amount of loss that has written down the tranche.

For example, let us assume that the bank is the buyer of protection on a 3% to 7% index tranche, Notional Reference Portfolio Size \$1.25 bn, Tranche Notional is 4%*\$1.25 bn = \$50 mm. The seller of protection is a hedge fund. Assume that a default takes place after one year on a payment date and the realised portfolio loss on the underlying reference portfolio increases from 2.5% to 3.5%. At this point the tranche is 'written down' since the realised loss is greater than the attachment point of 3%. The fractional tranche loss is:

(Realised Portfolio loss − Attachment point)/Tranche Width

= (3.5%−3%)/(7%−3%) = 12.5%.

The seller of protection makes a payment to the buyer of protection to compensate for this loss:

Notional*Tranche Width*Fractional Loss = \$1.25 bn*4%*12.5% = \$6.25 mm.

The buyer of protection makes future coupon payments to the seller of protection based on the written down notional of the tranche (or Tranche Notional Oustanding):

$$\text{Notional*Tranche Width*}(1 - \text{Fractional Loss}) = \$1.25\,\text{bn*4\%*}(1 - 0.125) = \$43.75\,\text{mm}$$

Assuming that the default occurred on a payment date (assuming no further defaults), the next quarterly coupon paid on the outstanding tranche notional is calculated as follows:

Written down tranche Notional*Coupon* Time to next quarterly coupon payment
= \$43.75 mm*275 bps* 0.25 = \$300,781.25

Impact of the 2007–8 credit crunch: new CDS contracts

One of the impacts of the 2007–8 financial market crisis was that CDS prices rose to hitherto unseen, astronomically high levels. The bankruptcy of Lehman Brothers (see Chapter 24) highlighted also the issue of counterparty risk for those market participants that had bought protection using CDS.

One response to this was that the markets changed the protocol for quoting CDS contracts. A new format for CDS in the North American market became effective from 8 April 2009. The new instruments traded with an upfront payment and fixed coupons of either 100 basis points for investment grade reference names or 500 basis points for sub-investment grade names.

The CDS 'Big Bang'

ISDA introduced a new supplement and protocol (the 'Big Bang' protocol) and a new standard North American corporate CDS contract with effect from April 2009 (SNAC). The ISDA supplement applied to new CDS transactions. It established credit determination committees, added auction settlement provisions and created backstop dates for credit and succession events. The Big Bang protocol applies to existing CDS transactions. The ISDA SNAC, also referred to as 100/500, applies to North American names denominated in any currency.

CDS and points upfront

For high-risk reference names, CDS spreads that have widened to a large extent are quoted by market makers with 'points upfront'. In this case, if a CDS trades with an upfront fee, a market counterparty buying protection must make an initial payment (a percentage of the notional value) as well as a running spread of 500 basis points.

The Bloomberg screen CDSW can be used to value CDS that are quoted with upfront fees, as shown at Figure 20.9. The trade example here is a five-year CDS quoted on 28 May 2009 on Virgin Media Finance plc. The CDS premium is 575 basis points, and if this traded in the US market there is now a 'points upfront' fixed fee to be paid on inception. This is shown on the screen as 2.885757%. Note

Figure 20.9 Bloomberg screen CDSW showing five-year CDS written on Virgin Media Finance plc name, traded 28 May 2009, with 'points upfront' valuation of 2.88%
© Bloomberg L.P. Used with permission. Visit www.bloomberg.com

also that the pricing model selected has been changed from the hitherto standard JPMorgan model to the 'ISDA standard upfront' model. Note also that in the field 'SNAC' the user has selected 'Y' for yes, indicating this contract is being traded in the US market and not in Europe.

Another response to the market that can be observed from Bloomberg screen CDSW concerns the recovery rate parameter. Previously this had defaulted to 40%. For a large number of lower-rated names this value has been tailor-set to levels ranging from 5% upwards.

Screen UPFR on Bloomberg shows the recovery rate for reference names now that the market no longer defaults automatically to 40%. Page 1 of this screen is shown at Figure 20.10. We see that recovery rates for this group of companies range from 15% to 40%.

General applications of credit derivatives

Credit derivatives have allowed market participants to separate and disaggregate credit risk, and thence to trade this risk in a secondary market (see, for example, Das, 2000). Initially portfolio managers used them to reduce credit exposure; subsequently they have been used in the management of portfolios, to enhance portfolio yields and in the structuring of synthetic CDOs. Banks use credit derivatives to transfer credit risk of their loan and other asset portfolios, and to take on credit exposure based on their views on the credit market. In this regard they also act as

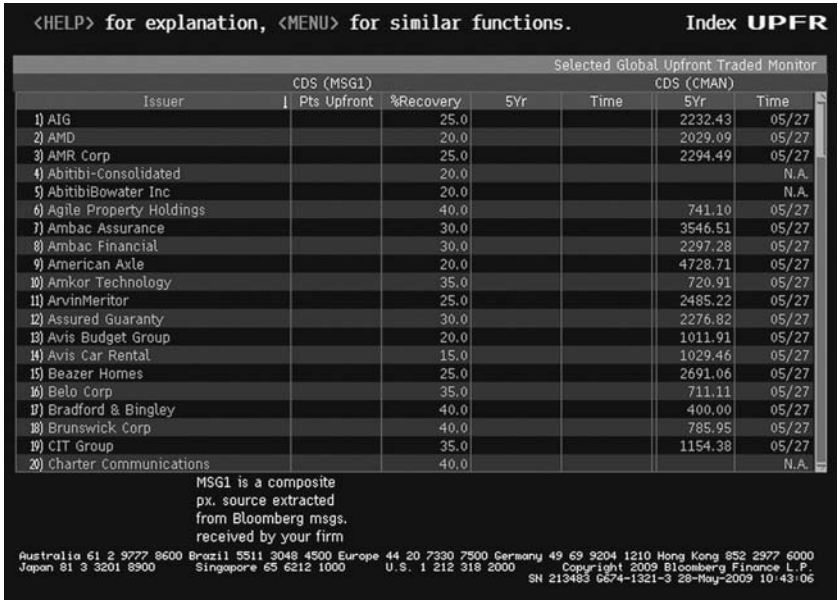


Figure 20.10 Bloomberg screen UPFR showing reference name recovery rates
© Bloomberg L.P. Used with permission. Visit www.bloomberg.com

credit derivatives market makers, running mismatched books in long and short-position CDSs and TRSs. This is exactly how they operate in the interest-rate market, using interest-rate swaps.

Use of credit derivatives by portfolio managers

Enhancing portfolio returns

Asset managers can derive premium income by trading credit exposures in the form of derivatives issued with synthetic structured notes. This would be part of a structured credit product. A pool of risky assets can be split into specific tranches of risk, with the most risky portion given the lowest credit rating in the structure. This is known as *multi-tranching*. The multi-tranching aspect of structured products enables specific credit exposures (credit spreads and outright default), and their expectations, to be sold to meet specific areas of demand. By using structured notes such as CLNs, tied to the assets in the reference pool of the portfolio manager, the trading of credit exposures is crystallised as added yield on the asset manager's fixed-income portfolio. In this way the portfolio manager enables other market participants to gain an exposure to the credit risk of a pool of assets but not to any other aspects of the portfolio, and without the need to hold the assets themselves.

Reducing credit exposure

Consider a portfolio manager who holds a large portfolio of bonds issued by a particular sector (say, utilities) and believes that spreads in this sector will widen in the

short term. Previously, in order to reduce her credit exposure she would have had to sell bonds; however, this might crystallise a mark-to-market loss and conflict with her long-term investment strategy. An alternative approach would be to enter into a CDS, purchasing protection for the short term; if spreads do widen these swaps will increase in value and may be sold at a profit in the secondary market. Alternatively the portfolio manager may enter into total return swaps on the desired credits. She pays the counterparty the total return on the reference assets, in return for Libor. This transfers the credit exposure of the bonds to the counterparty for the term of the swap, in return for the credit exposure of the counterparty.

Consider now the case of a portfolio manager wishing to mitigate credit risk from a growing portfolio (say, one that has just been launched). Figure 20.11 shows an example of an unhedged credit exposure to a hypothetical credit-risky portfolio. It illustrates the manager's expectation of credit risk building up to $250 million as assets are purchased, and then reducing to a more stable level as the credits become more established.[7] A three-year CDS entered into shortly afterwards provides protection on half of the notional exposure, shown as the broken line. The net exposure to credit events has been reduced by a significant margin.

Credit switches and zero-cost credit exposure

Protection buyers utilising CDSs must pay premium in return for laying off their credit risk exposure. An alternative approach for an asset manager involves the use of credit switches for specific sectors of the portfolio. In a credit switch the portfolio manager purchases credit protection on one reference asset or pool of assets, and simultaneously sells protection on another asset or pool of assets.[8]

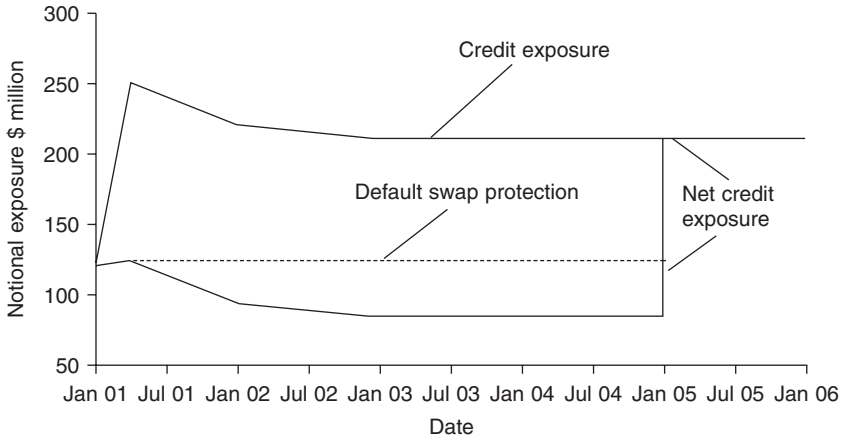


Figure 20.11 Reducing credit exposure

[7] For instance, the fund may be invested in new companies. As the names become more familiar to the market the credits become more 'established' because the perception of how much credit risk they represent falls.

[8] A pool of assets would be concentrated on one sector, such as utility company bonds.

So, for example, the portfolio manager would purchase protection for a particular fund and sell protection on another. Typically the entire transaction would be undertaken with one investment bank, which would price the structure so that the net cash flows would be zero. This has the effect of synthetically diversifying the credit exposure of the portfolio manager, enabling her to gain and/or reduce exposure to sectors as desired.

Exposure to market sectors

Investors can use credit derivatives to gain exposure to sectors for which they do not wish a cash market exposure. This can be achieved with an *index swap*, which is similar to a TRS, with one counterparty paying a total return that is linked to an external reference index. The other party pays a Libor-linked coupon or the total return of another index. Indices that are used might include the government bond index, a high-yield index or a technology stocks index. Assume that an investor believes that the bank loan market will outperform the mortgage-backed bond sector; to reflect this view he enters into an index swap in which he pays the total return of the mortgage index and receives the total return of the bank loan index.

Another possibility is synthetic exposure to foreign currency and money markets. Again we assume that an investor has a particular view on an emerging market currency. If he wishes, he can purchase a short-term (say one-year) domestic coupon-bearing note, whose principal redemption is linked to a currency factor. This factor is based on the ratio of the spot value of the foreign currency on issue of the note to the spot value on maturity. Such currency-linked notes can also be structured so that they provide an exposure to sovereign credit risk. The downside of currency-linked notes is that if the exchange rate goes the other way, the note will have a zero return, in effect a negative return once the investor's funding costs have been taken into account.

Trading credit spreads

Assume that an investor has negative views on a certain emerging-market government bond credit spread relative to UK gilts. The simplest way to reflect this view would be to go long a CDS on the sovereign, paying X bps. Assuming that the investor's view is correct and the sovereign bonds decrease in price as their credit spread widens, the premium payable on the credit swap will increase. The investor's swap can then be sold into the market at this higher premium.

Use of credit derivatives by banks

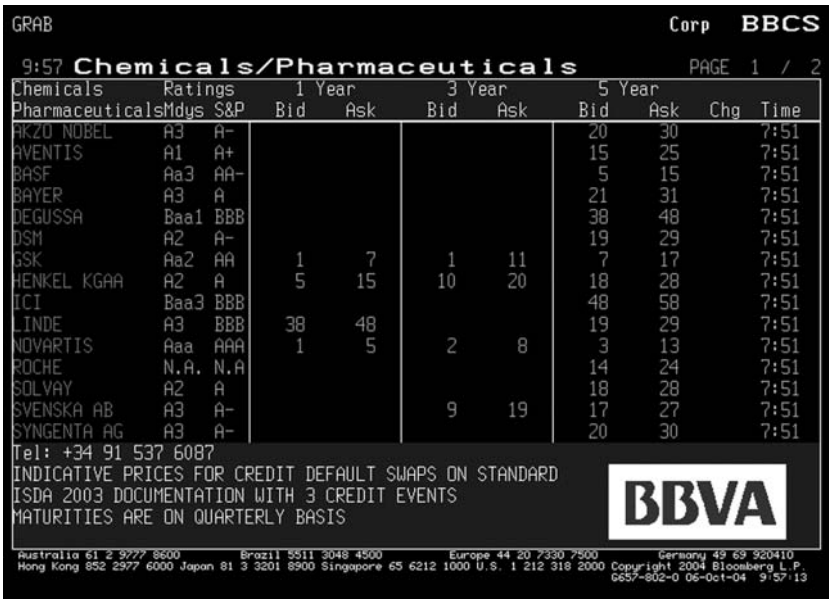
Banks use credit derivatives in exactly the same manner as portfolio managers – that is, in all the above we can replace 'fund managers' or 'investors' with 'banks'. But in fact banks were the first users of credit derivatives. The market developed as banks sought to protect themselves from loss due to default on portfolios of mainly illiquid assets, such as corporate loans and emerging-market syndicated loans. While securitisation was a well-used technique to move credit risk off the balance sheet, often this caused relationship problems with obligors, who would

feel that their close relationship with their banker was being compromised if the loans were sold off the bank's balance sheet. Banks would therefore buy protection on the loan book using CDSs, enabling them to hedge their credit exposure while maintaining banking relationships. The loan would be maintained on the balance sheet but would be fully protected by the CDSs.

To illustrate, consider Figure 20.12, which is a Bloomberg description page for a loan in the name of Haarman & Reimer, a chemicals company rated A3 by Moodys. We see that this loan pays 225 bps over Libor. Figure 20.13 shows the CDS prices page for A3-rated chemicals entities: Akzo Nobel is trading at 28 bps (to buy protection) as at 9 March 2004. A bank holding this loan can protect against default by purchasing this credit protection, and the relationship manager does not need to divulge this to the obligor. (In fact we may check the current price of this loan in the secondary market on the page BOAL, the Bank of America loan trading page on Bloomberg.)

The other major use by banks of credit derivatives is as a product offering for clients. The CDS market has developed exactly as the market did in interest-rate swaps, with banks offering two-way prices to customers and other banks as part of their product portfolio. Most commercial banks now offer this service, as they do in interest-rate swaps. In this role banks are both buyers and sellers of credit protection. Their net position will reflect their overall view on the market as well the other side of their customer business.

Figure 20.12 Haarman & Reimer loan description
© Bloomberg L.P. Used with permission. Visit www.bloomberg.com


GRAB Corp BBCS

9:57 Chemicals/Pharmaceuticals PAGE 1 / 2

Chemicals	Ratings		1 Year		3 Year		5 Year			
Pharmaceuticals	Mdys	S&P	Bid	Ask	Bid	Ask	Bid	Ask	Chg	Time
AKZO NOBEL	A3	A-					20	30		7:51
AVENTIS	A1	A+					15	25		7:51
BASF	Aa3	AA-					5	15		7:51
BAYER	A3	A					21	31		7:51
DEGUSSA	Baa1	BBB					38	48		7:51
DSM	A2	A-					19	29		7:51
GSK	Aa2	AA	1	7	1	11	7	17		7:51
HENKEL KGAA	A2	A	5	15	10	20	18	28		7:51
ICI	Baa3	BBB					48	58		7:51
LINDE	A3	BBB	38	48			19	29		7:51
NOVARTIS	Aaa	AAA	1	5	2	8	3	13		7:51
ROCHE	N.A.	N.A					14	24		7:51
SOLVAY	A2	A					18	28		7:51
SVENSKA AB	A3	A-			9	19	17	27		7:51
SYNGENTA AG	A3	A-					20	30		7:51

Tel: +34 91 537 6087
INDICATIVE PRICES FOR CREDIT DEFAULT SWAPS ON STANDARD ISDA 2003 DOCUMENTATION WITH 3 CREDIT EVENTS MATURITIES ARE ON QUARTERLY BASIS

BBVA

Australia 61 2 9777 8600 Brazil 5511 3048 4500 Europe 44 20 7330 7500 Germany 49 69 920410
Hong Kong 852 2977 6000 Japan 81 3 3201 8900 Singapore 65 6212 1000 U.S. 1 212 318 2000 Copyright 2004 Bloomberg L.P.
G657-802-0 06-Oct-04 9:57:13

Figure 20.13 Chemicals sector CDS prices from Banco Bilbao Vizcaya, 9 March 2004
©Bloomberg L.P. ©BBVA. Used with permission.

Example 20.2: Risks of synthetic positions and cash positions compared

Consider two investors in XYZ, one of whom owns bonds issued by XYZ while the other holds a credit-linked note referenced to XYZ. Following a deterioration in its debt situation, XYZ violates a number of covenants on its bank loans, but its bonds are unaffected. XYZ's bank accelerates the bank loan, but the bonds continue to trade at 85 cents on the dollar, coupons are paid and the bond is redeemed in full at maturity. However, the default swap underlying the CLN cites 'obligation acceleration' (of either bond or loan) as a credit event, so the holder of the CLN receives 85% of par in cash settlement and the CLN is terminated. However, the cash investor receives all the coupons and the par value of the bonds on maturity.

This example illustrates how, as CDSs are defined to pay out in the event of a very broad range of definitions of a 'credit event', portfolio managers may suffer losses as a result of occurrences that are not captured by one or more of the ratings agencies' rating of the reference asset. This results in a potentially greater risk for the portfolio manager than if it were actually to hold the underlying reference asset. Essentially, therefore, it is important for the range of definitions of a 'credit event' to be fully understood by counterparties, so that holders of default swaps are not taking on greater risk than is intended.

Credit derivatives pricing and valuation

Introduction

The pricing of credit derivatives should aim to provide a 'fair value' for the credit derivative instrument. In the sections below we discuss the pricing models currently used in the industry. The effective use of pricing models requires an understanding of the models' assumptions and the key pricing parameters, and a clear understanding of the limitations of a pricing model.

Issues to consider when carrying out credit derivative pricing include:

- implementation and selection of appropriate modelling techniques
- parameter estimation
- quality and quantity of data to support parameters and calibration
- calibration to market instruments for risky debt.

For credit derivative contracts in which the payout is on credit events other than default, the modelling of the credit evolutionary path is critical. If, however, a credit derivative contract does not pay out on intermediate stages between the current state and default, then the important factor is the probability of default from the current state.

Before continuing with this chapter, readers may wish to look at the section that discusses asset swap pricing methods, part of our discussion on the *basis*, in Chapter 7 of Choudhry (2004). This was commonly used at the inception of the credit derivatives market, but is rarely used today due to the inherent differences between asset swaps and other credit derivatives.

We now consider a number of pricing models as used in the credit derivative markets.

Pricing models

Pricing models for credit derivatives fall into two classes:

- Structural models
- Reduced form models.

We discuss these models below.

(i) Structural models

Structural models are characterised by modelling the firm's value in order to provide the probability of a firm default. The Black–Scholes–Merton option pricing framework is the foundation of the structural model approach. The default event is assumed to occur when the firm's assets fall below the book value of the debt.

Merton applied option pricing techniques to the valuation of corporate debt. By extension, the pricing of credit derivatives based on corporate debt may in some circumstances be treated as an option on debt (which is therefore analogous to an option on an option model).

Merton models have the following features:

- Default events occur predictably when a firm has insufficient assets to pay its debt.
- A firm's assets evolve randomly, and the probability of a firm default is determined using the Black–Scholes–Merton option pricing theory.

The constraint of structural models is that the behaviour of the value of assets and the parameters used to describe the process for the value of the firm's assets are not directly observable, and the method can be inconsistent with underlying market information for credit instruments.

(ii) Reduced form models

Reduced form models are a form of no-arbitrage model. These models can be fitted to the current term structure of risky bonds to generate no-arbitrage prices. In this way the pricing of credit derivatives using these models will be consistent with the market data on the credit-risky bonds traded in the market. These models allow the default process to be separated from the asset value, and are more commonly used to price credit derivatives.

Some key features of reduced form models include:

- Complete and arbitrage-free credit market conditions are assumed.
- Recovery rate is an input into the pricing model.
- Credit spread data are used to estimate the risk-neutral probabilities.
- Use of transition probabilities from credit agencies can be accommodated in some of these models. The formation of the risk-neutral transition matrix from the historical transition matrix is a key step.
- Default can take place randomly over time and the default probability can be determined using the risk-neutral transition matrix.

When implementing reduced form models it is necessary to consider issues such as the illiquidity of underlying credit-risky assets. Liquidity is often assumed to be present when we develop pricing models. However, in practice there may be problems when calibrating a model to illiquid positions, and in such cases the resulting pricing framework may be unstable and provide the user with spurious results. Another issue is the relevance of using historical credit transition data, in projecting future credit migration probabilities. In practice it is worthwhile reviewing the sensitivity of price to the historical credit transition data when using the model.

Reduced form models

Recent models which provide a detailed modeling of default risk include those presented by Jarrow, Lando and Turnbull (1997), Das and Tufano (1996) and Duffie and Singleton (1997). We consider these models in this section.

(i) The Jarrow, Lando and Turnbull (JLT) model

This model focuses on modeling default and credit migration. Its data and assumptions include the use of

- a statistical rating transition matrix which is based on historic data
- risky bond prices from the market used in the calibration process
- a constant recovery rate assumption. The recovery amount is assumed to be received at the maturity of the bond.
- a credit spread assumption for each rating level.

It also assumes no correlation between interest rates and credit rating migration.

The statistical transition matrix is adjusted by calibrating the expected risky bond values to the market values for risky bonds. The adjusted matrix is referred to as the risk-neutral transition matrix. The risk-neutral transition matrix is key to the pricing of several credit derivatives.

The JLT model allows the pricing of default swaps, as the risk-neutral transition matrix can be used to determine the probability of default. The model is sensitive to the level of the recovery rate assumption and the statistical rating matrix. It has a number of advantages; as the model is based on credit migration, it allows the pricing of derivatives for which the payout depends on such credit migration. In addition, the default probability can be explicitly determined and may be used in the pricing of credit default swaps.

The disadvantages of the model include the fact that it depends on the selected historical transition matrix. The applicability of this matrix to future periods needs to be considered carefully; for example, whether it adequately describes future credit migration patterns. In addition, it assumes all securities with the same credit rating have the same spread, which is restrictive. For this reason, the spread levels chosen in the model are a key assumption in the pricing model. Finally, the constant recovery rate is another practical constraint, as in practice the level of recovery will vary.

(ii) The Das–Tufano (DT) model

This model is an extension of the JLT model, and aims to produce the risk-neutral transition matrix in a similar way to the JLT model. However, the DT model uses stochastic recovery rates. The final risk-neutral transition matrix should be computed from the observable term structures. The stochastic recovery rates introduce more variability in the spread volatility. Spreads are a function of factors which may not be dependent solely on the rating level of the credit, as in practice credit spreads may change even though credit ratings have not changed. Therefore, to some extent the DT model introduces this additional variability into the risk-neutral transition matrix.

Various credit derivatives may be priced using this model; for example, credit default swaps, total return swaps and credit spread options. The pricing of these

products requires the generation of the appropriate credit-dependent cash flows at each node on a lattice of possible outcomes. The fair value may be determined by discounting the probability-weighted cash flows. The probability of the outcomes would be determined by reference to the risk-neutral transition matrix.

(iii) The Duffie–Singleton model

The Duffie–Singleton modeling approach considers the three components of risk for a credit-risky product; namely the risk-free rate, the hazard rate and the recovery rate.

The *hazard rate* characterises the instantaneous probability of default of the credit-risky underlying exposure. As each of the components above may not be static over time and a pricing model may assume a process for each of these components of risk, the process may be implemented using a lattice approach for each component. The constraint on the lattice formation is that this lattice framework should agree to the market pricing of credit-risky debt.

Here we demonstrate that the credit spread is related to risk of default (as represented by the hazard rate) and the level of recovery of the bond. We assume that a zero-coupon risky bond is maturing in a small time element Δt where:

λ is the annualised hazard rate
φ is the recovery value
r is the risk-free rate
s is the credit spread

and where its price P is given by

$$P = e^{-r\Delta t}((1 - \lambda\Delta t) + (\lambda\Delta t)\varphi) \tag{20.1}$$

Alternatively P may be expressed as:

$$P \cong e^{-\Delta t(r + \lambda(1 - \varphi)} \tag{20.2}$$

However, the usual form for a risky zero-coupon bond is

$$P = e^{-\Delta t(r + s)} \tag{20.3}$$

Therefore, we have shown that

$$s \cong \lambda(1 - \varphi) \tag{20.4}$$

This would imply that the credit spread is closely related to the hazard rate (that is, the likelihood of default) and the recovery rate.

This relationship between the credit spread, the hazard rate and recovery rate is intuitively appealing. The credit spread is perceived to be the extra yield (or return) an investor requires for credit risk assumed. For example:

- as the hazard rate (or instantaneous probability of default) rises, the credit spread increases; and
- as the recovery rate decreases the credit spread increases.

A 'hazard rate' function may be determined from the term structure of credit. The hazard rate function has its foundation in statistics, and may be linked to the instantaneous default probability.

The hazard rate function ($\lambda(s)$) can then be used to derive a probability function for the survival function $S(t)$:

$$S(t) = \exp^{-\int_0^t \lambda(s)ds} \tag{20.5}$$

The hazard rate function and survival probability may be determined by using the prices of risky bonds or credit default swaps. The lattice for the evolution of the hazard rate should be consistent with the hazard rate function implied from market data. An issue when performing this calibration is the volume of relevant data available for the credit.

Recovery rates

The recovery rate usually takes the form of the percentage of the par value of the security recovered by the investor.

The key elements of the recovery rate include:

- the level of the recovery rate
- the uncertainty of the recovery rate based on current conditions specific to the reference credit
- the time interval between default and the recovery value being realised.

Generally recovery rates are related to the seniority of the debt. Therefore the recovery value of the debt may change according to the seniority of the debt. Also, in practice, recovery rates exhibit significant volatility depending on the stage in the credit cycle.

Credit spread modelling

The examples that follow show how different models can sometimes lead to significantly different results.

Model 1: Lognormal process with volatility σ

$$dS(t) = \sigma S(t)dZ(t)$$

Model 2: Mean reversion normal model with jumps

$$dS(t) = \kappa(\mu - S(t))dt + \sigma dZ(t) + \gamma(dM(t) - \lambda dt)$$

where

$dS(t)$	is the change in the spread level over an element of time (dt)
dt	is the element of time over which the change in spread is modelled
$S(t)$	is the level of the credit spread at time t
κ	is the rate of mean reversion
μ	is the mean level of the spread
$dZ(t)$	is the Wiener increment $\varepsilon\sqrt{dt}$ where ε is N(0,1) (normal distribution mean 0 and standard deviation 1)
σ	is the volatility of the credit spread p.a.
λ	is the intensity of the Poisson process
γ	is the jump size of the credit spread
$dM(t)$	is the number of Poisson events

For example, let us compare the valuation results for a payer swaption, giving the holder the right to buy protection on a notional of \$10 mm for a CDS term of 5 years, with 6 months until option expiry (Table 20.1) and 12 months until option expiry (Table 20.2). We assume a flat interest-rate curve and spread curve, a spread at time 0 of 90 bps, and parameter values as shown in the tables. The valuation results are based on a simulation of 10,000 trials for Model 1 and Model 2.

The change in spread model from a lognormal process to a mean reversion normal model with jumps can significantly change pricing and risk results. We can see that the results are sensitive to the mean reversion level and the speed of reversion.

Table 20.1 Comparison of model results, expiry in six months

Notional \$10mm Payer swaption expiry in 6 months Risk-free rate = 4% Strike = 90 bps Normal volatility = 30 bps	*Model 1*	*Model 2*
$\kappa = 1$, $\mu = 90$ bps, $\lambda = 1$, $\gamma = 30$ bps	\$35,651	\$42,333
$\kappa = 1$, $\mu = 95$ bps, $\lambda = 1$, $\gamma = 30$ bps	\$36,246	\$44,758
$\kappa = 2$, $\mu = 95$ bps, $\lambda = 1$, $\gamma = 30$ bps	\$34,662	\$42,471
$\kappa = 2$, $\mu = 95$ bps, $\lambda = 2$, $\gamma = 50$ bps	\$36,314	\$53,421

Table 20.2 Comparison of model results, expiry in 12 months

Notional \$10 mm Payer swaption expiry in 12 months Risk-free rate = 4% Strike = 90 bps Normal volatility = 30 bps	*Model 1*	*Model 2*
$\kappa = 1$, $\mu = 90$ bps, $\lambda = 1$, $\gamma = 30$ bps	\$48,869	\$48,917
$\kappa = 1$, $\mu = 95$ bps, $\lambda = 1$, $\gamma = 30$ bps	\$49,286	\$53,841
$\kappa = 2$, $\mu = 95$ bps, $\lambda = 1$, $\gamma = 30$ bps	\$47,253	\$45,535
$\kappa = 2$, $\mu = 95$ bps, $\lambda = 2$, $\gamma = 50$ bps	\$50,751	\$57,540

This is an important point as the choice of inappropriate risk and pricing parameters can lead to unrealistic valuation and risk statistics. During the credit crunch there were cases when this occurred in practice, for example in CPDO (constant proportion debt obligations) product modelling, the issues relating to the choice and calibration of spread model had in some cases significantly underestimated the spread risk in the CPDO product.

Credit spread products

The forward credit spread

The forward credit spread can be determined by considering the spot prices for the risky security and risk-free benchmark security, while the forward yield can be derived from the forward price of these securities. The forward credit spread is the difference between the forward risky security yield and the forward yield on a risk-free security. The forward credit spread is calculated by using yields to the forward date and the yield to the maturity of the risky assets.

For example, the following data are used in determining the forward credit spread:

Current date	1 February 1998
Forward date	1 August 1998
Maturity	1 August 2006
Time period from current date to maturity	8 years and 6 months
Time period from current date to forward date	6 months
Yield to forward date:	
Risk-free security	6.25%
Risky security	6.50%
Yield to maturity:	
Risk-free security	7.80%
Risky security	8.20%

The forward yields (calculated from inputs above; see below for detailed derivation) are:

Risk-free security	7.8976%
Risky security	8.3071%

The details of the calculation of forward rates are as follows. For the risk-free security:

$$(1.0780)^{8\frac{6}{12}} = (1.0625)^{\frac{6}{12}} * (1 + rf_{riskfree})^{8}$$

where $rf_{riskfree}$ is the forward risk-free rate implied by the yields on a risk-free security. This equation implies that $rf_{riskfree}$ is 7.8976%. Similarly for the risky security we have:

$$(1.082)^{8\frac{6}{12}} = (1.0625)^{\frac{6}{12}} * (1 + rf_{riskfree})^{8}$$

where rf_{risky} is the forward risky rate implied by the yields on a risky security. This equation implies that rf_{risky} is 8.3071%.

Therefore the forward credit spread is the difference between the forward rate implied by the risky security less the forward rate implied by the yields on a risk-free security. In the example above, this is:

$$rf_{risky} - rf_{riskfree} = 8.3071 - 7.8976 = 0.4095\%$$

The current spread is equal to 8.20 − 7.80 = 0.40% = 40 bps. The difference between the forward credit spread and the current spread is 0.4095 − 0.40 = 0.0095% = 0.95 bps.

The calculation of the forward credit spread is critical to the valuation of credit spread products, as the payoff of spread forwards is highly sensitive to the implied forward credit spread.

Credit spread options

First-generation pricing models for credit spread options may use models as described in the section on spread models. The key market parameters in a spread option model include the forward credit spread and the volatility of the credit spread.

The volatility of the credit spread is a difficult parameter to determine and may be approached in different ways, including:

- the historical volatility of the difference between the reference asset yield and the yield on a risk-free benchmark
- estimation of the historical volatility by considering the components (historic volatility of the reference asset yield, historic volatility of the benchmark yield, correlation of the returns between the reference asset yield and the benchmark yield)
- estimation of the volatility of the spread by using the implied volatility of the reference asset yield, implied volatility of the benchmark yield and a suitable forward-looking estimate of the correlation between the returns on the reference asset yield and benchmark asset yield.

If the model incorporates mean reversion, then other key inputs will include the mean reversion level and the rate of mean reversion. These inputs cannot be observed directly, and the choice should be supported by the model developers and constantly reviewed to ensure that they remain relevant. Other inputs include:

- the strike price
- the time to expiry
- the risk-free rate for discounting.

A key issue with credit spread options is ensuring that the pricing models used will calibrate to the market prices of credit-risky reference assets. The recovery of forward prices of the reference asset would be a constraint to the evolution of the credit spread. More complex spread models may allow for the correlation between

the level of the credit spread and the interest-rate level. The reduced form models described earlier are a new generation of credit derivative pricing models which are now increasingly being used to price spread options.

Asset swaps

Assume that an investor holds a bond and enters into an asset swap with a bank. Then the value of an asset swap is the spread the bank pays over or under Libor. This is based on the following components:

- value of the coupons of the underlying asset compared to the market swap rate
- the accrued interest and the clean price premium or discount compared to par value. Thus when pricing the asset swap it is necessary to compare the par value to the underlying bond price.

The spread above or below Libor reflects the credit spread difference between the bond and the swap rate.

For example, let us assume that we have a credit-risky bond with the following details:

Currency:	EUR
Issue date:	31 March 2000
Maturity:	31 March 2007
Coupon:	5.5% per annum
Price (dirty):	105.3%
Price (clean):	101.2%
Yield:	5%
Accrued interest:	4.1%
Rating:	A1

To buy this bond, the investor would pay 105.3% of par value. The investor would receive the fixed coupons of 5.5% of par value. Let us assume that the swap rate is 5%. The investor in this bond enters into an asset swap with a bank in which the investor pays the fixed coupon and receives Libor +/− spread.

The asset swap price (that is, the spread) on this bond has the following components:

- The value of the excess value of the fixed coupons over the market swap rate is paid to the investor. Let us assume that in this case this is approximately 0.5% when spread into payments over the life of the asset swap.
- The difference between the bond price and par value is another factor in the pricing of an asset swap. In this case the price premium which is expressed in present value terms should be spread over the term of the swap and treated as a payment by the investor to the bank. (If a dirty price is at a discount to the par value, then the payment is made from the bank to the investor.) For example, in this case let us assume that this results in a payment from the investor to the bank of approximately 0.23% when spread over the term of the swap.

These two elements result in a net spread of 0.5% − 0.23% = 0.27%. Therefore, the asset swap would be quoted as Libor + 0.27% (or Libor plus 27 bps).

Total return swap (TRS) pricing

The present value of the two legs of the TRS should be equivalent. This would imply that the level of the spread is therefore dependent on the following factors:

- credit quality of the underlying asset
- credit quality of the TRS counterparty
- capital costs and target profit margins
- funding costs of the TRS provider, as it will hedge the swap by holding the position in the underlying asset.

The fair value for the TRS will be the value of the spread for which the present value of the Libor +/– spread leg equals the present value of the returns on the underlying reference asset. The present value of the returns on the underlying reference asset may be determined by evolving the underlying reference asset. The expected value of the TRS payoff at maturity should be discounted to the valuation date.

The reduced form models described earlier are a new generation of credit derivative pricing models which are now increasingly being used to price total return swaps.

Forward starting CDS

The forward starting CDS is a key building block for products that reference the future level of the CDS spread. We determine the forward CDS spread by considering the term structure of CDS spreads and the risky annuity level for the relevant maturities.

Let

- $F_{t,\ T0,\ T}$ be the forward CDS spread at time t, for a CDS contract starting at time T_0 and maturing at time T
- $S_{t,T}$ be the CDS spread at time t for a maturity T
- $RPV01\ (t,T)$ is the value of the risky annuity of \$1 p.a. payable from time t to T. The credit risk is based on the survival curve associated with the credit.

The rationale for determining $F_{t,\ T0,\ T0+T}$ is to equate the cost of credit protection over the period t to T and the cost of the credit protection from t to T_0 plus the forward CDS protection from T_0 to T:

PV of Premium leg of cover credit protection from t to T = *PV* of Premium leg of cover credit protection from t to T_0 plus *PV* of Premium leg of forward credit protection from T_0 to T:

$$S_{t,T} * RPV01(t,T) = S_{t,T0} * RPV01(t,T) + [F_{t,T0,T} * (RPV01(t,T) - RPV01(t,T_0))]$$

So the estimate of the forward CDS $F_{t,\ T0,\ T}$ is:

$$F_{t,T0,T} = [S_{t,T} * RPV01(t,T) - S_{t,T0} * RPV01(t,T_0)]/[RPV01(t,T) - RPV01(t,T_0)]$$

For example, assume that we wish to determine the forward credit spread today for protection over the period three to seven years in the future. Let us also assume that we have available the following data:

- current CDS spread for a three-year period starting at time 0, $S_{0,3} = 50$ bps
- RPV01 (0,3) = 2.75 risky annuity at time 0 for a three-year period
- current CDS spread for a seven-year period starting at time 0, $S_{0,7} = 70$ bps
- RPV01 (0,7) = 6.5 risky annuity at time 0 for a seven-year period.

The forward starting CDS at time 0 for the period beginning in three years and ending in seven years $F_{0,3,7}$ is:

$$F_{0,3,7} = [(70\,bps * 6.5) - (50\,bps * 2.75)]/[6.5 - 2.75] = 84.67\,bps$$

Credit default swaption (CDS swaption)

The CDS swaption gives the holder the right but not the obligation to enter into a single name CDS contract at expiry date. Investors can use options to take views on credit spread volatility and also make directional trades to either hedge or speculate.

The payer swaption gives the option holder the right to buy protection at expiry at the agreed strike level. It gives the option holder the right to go short the credit. The holder will exercise the option if the CDS spread level at the expiry date is higher than the strike level. The scenario in which spread levels are higher than the strike level would make the option valuable to the holder, since it is similar to having the right to sell a credit-risky bond at a higher price than market price at expiry. Payers are similar to put options on bonds, as they give investors the right to sell credit risk.

The receiver swaption gives the option holder the right to sell protection at the expiry date, at the agreed strike level. It gives the option holder the right to go long the credit. The holder will exercise the option if the CDS spread level at the expiry date is lower than the strike level. The scenario in which the spread levels are lower than the strike level would make the option valuable to the holder, since it is similar to having the right to buy a credit-risky bond at a lower price than market price at expiry. Receivers are similar to bond call options, as they give investors the right to buy credit risk.

The option contract can be structured to knock out if there is a credit event before expiry date with no cost – or it may be structured not to knock out before expiry date.

The valuation of the knockout version of the contract on a single name CDS can be considered by using an equation that is similar to the Black (1976) bond option pricing model:

$$\text{Value of a payer swaption (knockout)} = \textit{Notional Value} * [RPV01(0,T_E) - RPV\,01(0,T)] * [F_{0,TE,T} * \Phi(d_1) - K * \Phi(d_2)]$$

where

- $d_1 = [LN(F_{0,TE,T}/K) + (\sigma^2 * T_E)/2]/[\sigma * \sqrt{(T_E)}]$
- $d_2 = d_1 - \sigma * \sqrt{(T_E)}$
- $[RPV01(0,T_E) - RPV01(0,T)]$ is equal to the value of the risky forward annuity from T_E to T.

The model dynamics for the evolution of the spread are an important assumption; the choice of model is important and can lead to different values. The market standard for the CDS swaption is based on the Black (1976) model.

Credit index options

Credit index options may be either:

- a payer swaption (giving the holder the right to buy protection at the strike level *K* at the expiry date) or
- a receiver swaption (giving the holder the right to sell protection at the strike level *K* at the expiry date).

The option typically references an underlying credit index. Credit index options do not knock out if there are credit events before expiry date.

For example, at the expiry date of a payer swaption, the holder will determine if they wish to exercise the option and buy protection. In order to make this exercise decision the holder will consider the two areas that affect the value on exercising the option:

1. The receipt of the value of payments for any credit events on the credit index
2. the value of entering into a transaction to buy protection on the credit index (the 'off the run' index if there have been credit events or the 'on the run' index if there are no credit events) at the agreed strike level *K*. The value of this transaction can be analysed to consist of two elements:
   (i) the present value of the difference between the fixed coupon on the index and strike level *K*, discounted using a risky spread equal to the strike level
   (ii) the present value of the difference between the index spread level at expiry and the coupon on the index, discounted using a risky spread equal to the spread level at expiry. This is the value of the short index position at expiry.

The exercise decision can only be made after considering the value of these elements. If the total value is greater than zero, the holder should exercise the payer swaption. The settlement mechanics are designed to allow for the value of 1 and 2, to settle at the expiry date.

The pricing of credit index options requires an adjustment to the Black (1976) model that is used for single name CDS options, the adjustment is made to the forward and the strike inputs, to allow for the non-knockout features of the credit index option contract.

Credit curves

The *credit curves* (or *default swap curves*) reflect the term structure of spreads by maturity (or *tenor*) in the credit default swap markets. The shape of the credit curves is influenced by the demand and supply for credit protection in the CDS market, and reflects the credit quality of the reference entities (both specific and systematic risk). The changing levels of credit curves provide traders and arbitrageurs with the opportunity to measure relative value and establish credit positions.

In this way, any changes of shape and perceptions of the premium for CDS protection are reflected in the spreads observed in the market. In periods of extreme price volatility, as seen in the middle of 2002, the curves may invert to reflect the fact that the cost of protection for shorter-dated protection trades at wider levels than the longer-dated protection. This is consistent with the pricing theory for credit default swaps.

The probability of survival for a credit may be viewed as a decreasing function against time. The survival probabilities for each traded reference credit can be derived from its credit curve. The survival probability is a decreasing function, because it reflects the fact that the probability of survival for a credit reduces over time – for example, the probability of survival to year 3 is higher than the probability of survival to year 5.

Under non-volatile market conditions, the shape of the survival probability and the resulting credit curve will take a different form from the shape implied in volatile market conditions; the graphs may change to reflect the higher perceived likelihood of default. For example, the shape of the survival probability may take the form shown in Figure 20.14. The corresponding credit curves consistent with

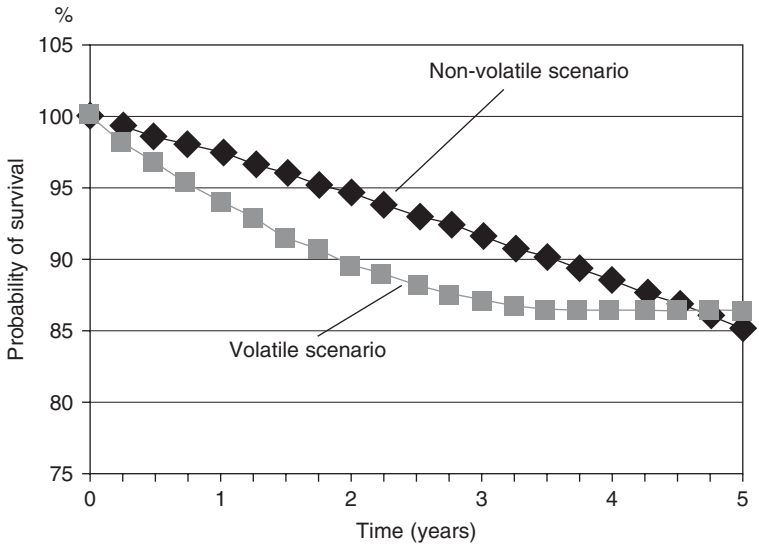


Figure 20.14 Probability of survival

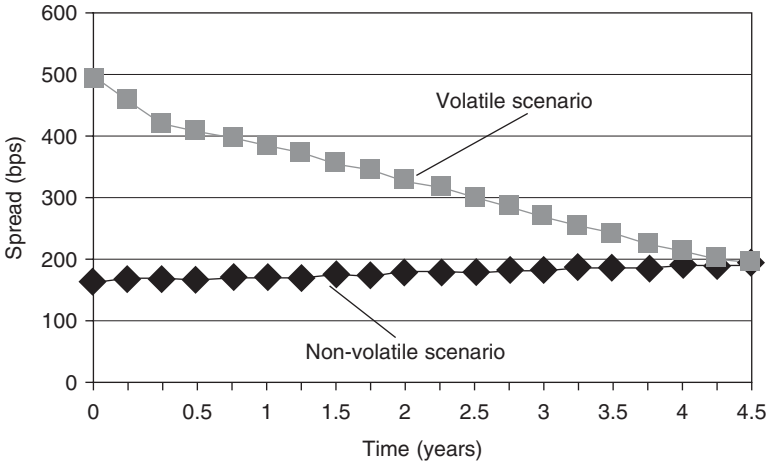


Figure 20.15 Credit curves

these survival probabilities take the form shown in Figure 20.15. This shows that the credit curve inversion is consistent with the changes in the survival probability functions. In this analysis, we assume that the assumed recovery rate for the 'cheapest to deliver' bond remains the same at 35% of notional value.

Appendix 20.1: Sample term sheet for a credit default swap traded by XYZ Bank plc

Draft Terms – Credit Default Swap

1. General Terms

Trade Date	Aug 5, 2003
Effective Date	Aug 6, 2003
Scheduled Termination Date	Jul 30, 2005
Floating Rate Payer ('Seller')	XYZ Bank plc, London branch
Fixed Rate Payer ('Buyer')	ABC Investment Bank plc
Calculation Agent	Seller
Calculation Agent City	New York
Business Day	New York
Business Day Convention	Following
Reference Entity	Jackfruit Records Corporation
Reference Obligation	Primary Obligor: Jackfruit Records
Maturity	Jun 30, 2020
Coupon	0%
CUSIP/ISIN	xxxxx
Original Issue Amount	USD 100,000,000
Reference Price	100%
All Guarantees	Not Applicable

2. Fixed Payments

Fixed Rate Payer Calculation Amount	USD 7,000,000
Fixed Rate	0.3% per annum
Fixed Rate Payer Payment Date(s)	Oct 30, Jan 30, Apr 30, Jul 30, starting Oct 30, 2003
Fixed Rate Day Count Fraction	Actual/360

3. Floating Payments

Floating Rate Payer Calculation Amount	USD 7,000,000
Conditions to Payment	Credit Event Notice (Notifying Parties: Buyer or Seller) Notice of Publicly Available Information: Applicable (Public Source: Standard Public Sources. Specified Number: Two)
Credit Events	Bankruptcy Failure to Pay (Grace Period Extension: Not Applicable. Payment Requirement: $1,000,000)
Obligation(s)	Borrowed Money

4. Settlement Terms

Settlement Method	Physical Settlement
Settlement Currency	The currency in which the Floating RatePayer Calculation Amount is denominated

TERMS RELATING TO PHYSICAL SETTLEMENT

Physical Settlement Period	The longest of the number of business days for settlement in accordance with the then-current market practice of any Deliverable Obligation being Delivered in the Portfolio, as determined by the Calculation Agent, after consultation with the parties, but in no event shall be more than 30 days
Portfolio	Exclude Accrued Interest
Deliverable Obligations	Bond or Loan
Deliverable Obligation Characteristics	Not Subordinated Specified Currency – Standard Specified Currencies Maximum Maturity: 30 years Not Contingent Not Bearer Transferable Assignable Loan Consent Required Loan
Restructuring Maturity Limitation	Not Applicable
Partial Cash Settlement of Loans	Not Applicable
Partial Cash Settlement of Assignable Loans	Not Applicable
Escrow	Applicable

5. Documentation

Confirmation to be prepared by the Seller and agreed to by the Buyer. The definitions and provisions contained in the 2003 ISDA *Credit Derivatives Definitions*, as published by the International Swaps and Derivatives Association, Inc., as supplemented by the May 2003 Supplement, to the *2003 ISDA Credit Derivatives Definitions* (together, the 'Credit Derivatives Definitions'), are incorporated into the Confirmation.

6. Notice and Account Details

Telephone, Telex and/or Facsimile Numbers and Contact Details for Notices	Buyer: Phone: Fax: Seller: A.N. Other

	Phone: +1 212-xxx-xxxx
	Fax: +1 212-xxx-xxxx
Account Details of Seller	84-7512562-85

Risks and Characteristics

Credit Risk. An investor's ability to collect any premium will depend on the ability of XYZ Bank plc to pay.

Non-Marketability. Swaps are not registered instruments and they do not trade on any exchange. It may be impossible for the transactor in a swap to transfer the obligations under the swap to another holder. Swaps are customised instruments and there is no central source to obtain prices from other dealers.

References and bibliography

Choudhry, M. *The Bond and Money Markets: Strategy, Trading, Analysis*, Butterworth-Heinemann, 2001.

Choudhry, M. 'Some issues in the asset-swap pricing of credit default swaps', in F. Fabozzi (ed.), *Professional Perspectives on Fixed Income Portfolio Management*, vol. 4, Wiley, 2003.

Choudhry, M. *Structured Credit Products: Credit Derivatives and Synthetic Securitisation,* 2nd edition Wiley Asia, 2010.

Das, S. *Credit Derivatives and Credit-Linked Notes*, 2nd edition, Wiley, 2000.

Das, S. and Tufano, P. 'Pricing credit sensitive debt when interest rates, credit ratings and credit spreads are stochastic', *Journal of Financial Engineering*, 5, 1996, pp. 161–98.

Das, S. and Tufano, P. 'Pricing credit-sensitive debt when interest rate, credit ratings and credit spreads are stochastic', *Journal of Finance Engineering*, 7, 1996, pp. 133–49.

Duffie, D. 'Credit swap valuation', *Financial Analysts Journal*, Jan.–Feb. 1999, pp. 73–87.

Duffie, D. and Huang, M. 'Swap rates and credit quality', *Journal of Finance*, 51, 3, 1996.

Duffie, D. and Singleton, K. 'Modelling term structures of defaultable bonds', *Review of Financial Studies,* 11, 1997, pp. 61–74.

Duffie, D. and Singleton, K. 'Modelling term structure models of defaultable bonds', *Review of Financial Studies*, 12, 1999, pp. 687–720.

Francis, J., Frost, J. and Whittaker, G. *Handbook of Credit Derivatives*, New York: Irwin, 1999.

Hogg, R. and Craig, A. *Introduction to Mathematical Statistics*, 3rd edn, Macmillan, 1970.

Hull, J. and White, A. 'Valuing credit default swaps I: no counter-party default risk', *Journal of Derivatives*, 8, 1, 2000.

Jarrow, R., Lando, D. and Turnbull, S. M. 'A Markov model for the term structure of credit spreads', *Review of Financial Studies*, 10, 1997, pp. 481–523.

Jarrow, R. A. and Turnbull, S. M. 'Pricing options on derivative securities subject to credit risk', *Journal of Finance*, 50, 1995, pp. 53–8.

Longstaff, F. A. and Schwartz, E. S. 'Valuing credit derivatives', *Journal of Fixed Income*, 5, 1, 1995, pp. 6–12.

Merton, R. C. 'On the pricing of corporate debt: the risk structure of interest rates', *Journal of Finance*, 29, 1974, pp. 449–70.

Tolk, J. 'Understanding the risks in credit default swaps', Moodys Investors Service Special Report, 16 March, 2001.

Yekutieli, I. 'With bond stripping, the curve's the thing', Internal Bloomberg report, 1999.

This page intentionally left blank

PART V
Equity Capital Markets

In Part V of the book we present an overview of the basic concepts in equity valuation and analysis. Analysis of the equity instrument is different from that of the debt instrument, for a number of fundamental reasons. These are touched on at the start of Chapter 21. We then discuss the basic concepts of the dividend valuation model and dividend growth model. The second chapter in Part V looks at fundamentals of ratio analysis and the cost of capital.

This page intentionally left blank

21
Introduction to Equity Instrument Analysis

Equity instruments call for a different approach in their analysis compared with debt securities. Compared with conventional bonds, shares do not pay a fixed cash flow during their life and do not have a fixed maturity date. Instead, the future cash flows of shares cannot be determined with certainty and must be assumed, and as they are in effect perpetual securities they have no redemption value. In addition they represent a higher form of risk for their holders. The analysis and valuation of shares therefore call for different techniques from that of bonds. In addition the interests of shareholders and bondholders sometimes often sit on opposite ends of the risk spectrum. While a firm's bondholders are to some extent primarily concerned with financial probity and the maintenance (or upgrade) of its credit rating, shareholders gain, at least in the short term, from high-risk and high-return strategies where favourable perceptions lead to a short-term rise in the share price.

In this chapter we present basic concepts in investment analysis for equities, beginning with the financial structure of the firm. We then consider the fair pricing of an equity, and dividend policy. In the following chapter we introduce ratio analysis and basic concepts in assessing the cost of capital.

Firm financial structure and company accounts

A corporate entity or *firm* is governed by the types of equity capital it can issue as stipulated in its *memorandum* and *articles of association*. In the past, in the UK market at least, firms would issue different classes of shares including 'A' shares that carried restricted voting rights. However this was not encouraged by the Stock Exchange and the most common form of share in the market is the *ordinary share*, which is known as *common stock* in the US market. The holders of ordinary shares are entitled to certain privileges, including the right to vote in the running of the company, the right to dividend payments and the right to subscribe for further shares ahead of non-shareholders, in the event of a new issue.

Dividends are only payable after liabilities to all other parties with a claim on the company, including bondholders, have been discharged. Shares are issued with a par value, but this has no relevance to their analysis and is frequently for a token amount such as £0.10 or £0.25. Firms also issue *preference shares* which are a type

of hybrid between shares and bonds, but these are less common and we will ignore them in this chapter. An introduction to preference shares is given in Chapter 23 of Choudhry (2001).

We begin by considering the financial structure of the firm, which traditionally was of vital importance to shareholders. We say 'traditionally' because many of the accepted tenets of fundamental investment analysis were in effect ignored by investors during the so-called 'dot.com' bull market of the late 1990s. However as we write, the various technology indices have been falling dramatically, and it is perhaps a good time to reacquaint ourselves with the basic concepts.

We can consider the importance to shareholders of the financial structure of a firm by comparing the interests of shareholders with those of bondholders. Unlike shareholders, bondholders have a prior contractual claim on the firm. This means that as and when the contractual claim is covered, bondholders have no further interest in the firm. Put another way, as long as the firm is able to meet its contractual commitments, which are interest and principal payments owing to creditors, bondholders will be satisfied.[1] On the other hand, shareholders have what is known as a *residual claim* on the firm. As the owners of the firm, they will be concerned about the overall value of the firm and that this is being maximised. Hence they are (in theory) keenly concerned with the financial structure of their firm, as well as its long-term prospects.

We begin, therefore, with a review of company accounts. Firms are required by law to produce accounts, originally under the belief that owners should be kept informed about how the directors are managing the company. In the UK for example this is stipulated in section 226 of the Companies Act 1985, which updated previous versions of the Act.

The balance sheet

The balance sheet is a snapshot in time of the asset value of a company. You should be familiar with the two sources of corporate financing in a developed economy: debt finance sourced from lenders including banks, finance houses and directly from the market through bond issues, and equity finance sourced from shareholders and retained profits. Put simply, once a corporate entity has repaid all its debt financing, the remaining funds are the property of the shareholders. Hence we may state:

$$\text{assets} - \text{liabilities} = \text{shareholders' funds}.$$

Again in simple terms, the valuation of one share in the company is a function of the total assets of the company less the liabilities. So as a firm's assets decrease in value, shareholders will experience a decrease in their share value, while the opposite occurs if there is an increase in firm assets. This explains why a corporate balance sheet always balances.

[1] This is perhaps too simplistic. Bondholders will also be concerned if any developments affect the perceived ability of the firm to meet its future liabilities, such as a change of credit rating. Such events will affect the value of the bonds issued by the firm, which is why they will be of concern to bondholders.

A company balance sheet may be put together using one of three different approaches: *historic cost book value, current cost* or *market value*. Equity analysts' preference is for the market value basis, which records the value of assets and liabilities in the balance sheet using current market values. For liabilities this is relatively straightforward to undertake if the firm is listed on an exchange and there is a liquid market in its shares; the net value of the firm can be taken to be the difference between the market value of the firm's ordinary shares and the *book value* of the shares. The latter is the par value of the shares plus the share premium and accumulated retained earnings. It is more problematic to determine a market value for firm assets, however. For instance, what is the market value of two-year-old photocopying machines? In fact the majority of incorporated institutions do not have their shares traded on an exchange, and so a market value balance sheet is rarely released.

The most common balance sheet approach uses the historic cost book value approach, in which assets and liabilities are valued at their original cost, known as historic cost. The net worth of the company is calculated as the sum of the share capital and retained profits (*reserves*). It is rare to observe balance sheets presented using the current cost book value approach, which values assets at current replacement cost.

A hypothetical company balance sheet in shown in Table 21.1.

Note that the balance sheet orders assets and liabilities in terms of their maturity. Fixed assets are recorded first, followed by current assets less current liabilities. This value, the *net current assets*, indicates whether the company is able to cover its short-term liabilities with its current assets. The net current asset value is added to the fixed asset value, resulting in the value of the firm's assets less current liabilities.

The balance sheet then records long-term liabilities, and these are subtracted from the previous figure, showing the total value of the company once all liabilities have been discharged. This is also known as *shareholders' funds*, and represents (in theory) the amount that would be distributed to them in the event that the firm was wound up at this point.

Shareholders' funds are represented by the capital and reserves entries. *Share capital* is the sum of the *issued share par value* and the *share premium*. These are defined as follows:

- *Paid-up share capital:* the nominal value of the shares, which represents the total liabilities of the shareholders in the event of winding up, and which has been paid by shareholders.
- *Share premium:* the difference between market value of the shares and the nominal or par value.

The entry for 'profit and loss account' sometimes appears as *retained earnings*. This is the accumulated profit over the life of the company that has not been paid out as dividend to shareholders, but has been reinvested in the company. The profit and loss account is part of the firm's reserves, and its calculation is arrived at via a separate financial statement.

Table 21.1 Hypothetical corporate balance sheet

Balance sheet as at 31 December 2001		
	£m	£m
Fixed assets		675
Long-term investments		98
		773
Current assets		
Stock	365	
Debtors	523	
Cash	18	
	906	
Short-term liabilities		
Creditors	355	
Short-dated loans	109	
Bank overdraft	88	
Corporation tax	91	
Planned dividend	66	
	709	
Net current assets		197
Total assets less current liabilities		970
Liabilities falling due after 12 months		
Creditors	28	
Long-dated debt, bonds	400	(428)
Net assets		542
Share capital		
Ordinary shares issued		170
Preference shares		30
		200
Capital and Reserves		
Paid-up capital	25	
Share premium	109	
Profit and loss account (reserves)	208	342
Shareholders' funds		542

Fixed assets include items such as factory buildings, property holdings, etc
Short-term liabilities are those falling due within 12 months

The profit and loss account

The profit and loss account, also known as the *income statement*, shows the profit generated by a firm, separating out the amount paid to shareholders and that retained in the company. Hence the profit and loss account is also a statement of retained earnings. Unlike the balance sheet, which is a snapshot in time, the income statement is a rolling total of retained profit from the last accounting period to the current one. Generally this period is one year.

The calculation of the profit and loss account[2] is relatively straightforward, recording income less expenses. A firm's income is that generated from its business

[2] Strictly speaking, it is a profit *or* loss account, as the firm would have made either a profit or loss in the accounting period.

activities, and so excludes share capital or loan funding. The expenses are daily costs of running the business, and so exclude items such as plant and machinery which are considered 'capital' expenditure and recorded as fixed assets in the balance sheet. Due to the different accounting conventions and bases in use, it is possible for two identical companies to produce very different profit and loss statements. This is a complex and vast subject, well outside the scope of this book, so we will not enter it in any more depth. A good overview of accounting principles in the context of corporate finance is given in Higson (1995).

A hypothetical profit and loss statement is shown in Table 21.2.

In the context of a profit and loss statement, the *net profit* is the gross profit minus business operating expenses. This is an accurate measure of the profit that the firm's managers have generated. The more efficiently managers run the business, the lower its expenses will be, and correspondingly the higher the net profit will be. Tax expenses are outside the control of the firm's managers and so appear afterwards. Extraordinary items are deemed to be those generating income that are outside the ordinary business activities of the company, and expected to be one-off or rare occurrences. This might include the disposal of a subsidiary, for example.

Consolidated accounts

Consolidated accounts are produced when a company has one or more subsidiaries; the accounts of the individual undertakings are combined into a single consolidated account for shareholders. In the UK this is required under the Companies Act 1985, based on the belief that a company's business will be closely linked to that of any subsidiary that it owns, and therefore its shareholders require financial statements on the combined entity. At the same time the subsidiaries also produce their own balance sheets and profit and loss accounts.

Table 21.2 Hypothetical corporate profit and loss statement

	£m	*£m*	*£m*
Operating revenue	737		
Operating costs	(389)		
Gross operating profit		348	
Expenses			
Administration	(19)		
Sales	(67)		
Financial	(27)		
		(113)	
Net profit		235	
Taxation		(78)	
Profit on ordinary activities after tax		157	
Extraordinary items		—	
		157	
Dividends			30
Retained profit			127
Retained profit brought forward			81
Retained profit carried forward			208

Valuation of shares

In this section we present some fundamental concepts in equity valuation. The approaches described seek to determine a share's *fair value*, which can then be compared with its market value. We consider first dividend valuation methods and then the *expected earnings* method.

Dividend valuation model

We assume a corporate entity that pays an annual dividend. An investor thinking of purchasing a number of shares of this company, and holding them for one year, will expect to receive the annual dividend payment during the time he holds the shares, as well as the share proceeds on disposal. The fair value that the investor would be prepared to pay today for the shares is given by

$$P_t = \frac{E(DV_{t+1})}{1+r} + \frac{E(P_{t+1})}{1+r} \tag{21.1}$$

where

P_t	is the price of the shares at time t
r	is the required rate of return
$E(DV_t + 1)$	is the expected annual dividend one year after t
$E(P_t + 1)$	is the expected price of the share one year after t.

The rate of return r can be related to the company's cost of capital plus a spread, or some other market-determined discount rate. The return generated by the shareholding is split between the income element, given by the dividend amount $(DV_t + 1)$ and the capital gain element given by $(P_{t+1} - P_t)$.

Following (21.1) we may say that

$$E(P_{t+1}) = \frac{E(DV_{t+2})}{1+r} + \frac{E(P_{t+2})}{1+r} \tag{21.2}$$

and by using substitution and continuing for successive years we obtain

$$P_0 = \sum_{t=1}^{T} \frac{E(DV_t)}{(1+r)^t} + \frac{E(P_T)}{(1+r)^T} \tag{21.3}$$

for the period beginning now (at $t = 0$) and where DV_t is the dividend per share in year t. If we extend T to the limiting factor as it becomes close to infinity, the share price element will disappear, so (21.3) transforms to:

$$P_0 = \sum_{t=1}^{\infty} \frac{E(DV_t)}{(1+r)^t} \tag{21.4}$$

Expression (21.4) is the dividend discount model, and one can observe its similarity to the bond price/redemption yield expression straight away. However the comparison of share and bond values using this method is fraught with complications, and so it is rarely undertaken. This reflects problems with the approach. First, the model assumes a constant discount rate or cost of capital over time, which is unrealistic; for this reason the appropriate *t*-period zero-coupon interest rate is sometimes used as the discount rate, with a spread added to the government rate to reflect additional risk associated with holding the share. Other problems associated with the model include:

- problems of divergence associated with the infinite time period implied by (21.4)
- an expectation of infinite dividend payments.

These issues may be resolved in practice by introducing further assumptions, considered in the next section as part of corporate dividend policy.

Dividend growth model

Following the dividend valuation model, let us assume a constant growth rate in dividend payments given by c. The dividend valuation model is then given by

$$P_0 = \sum_{t=0}^{\infty} \frac{E(DV_t)}{(1+r)^t} = \sum_{t=0}^{\infty} \frac{DV_0(1+c)^t}{(1+r)^t} \tag{21.5}$$

Assuming further that $r > c$, then it can be shown[3] that the sum of this infinite series is given by

$$P_0 = \frac{DV_1}{r-c} \tag{21.6}$$

What (21.6) states is that, under a constant dividend growth rate c that is less than the required rate of return r, the value of a company's share is the year 1 dividend divided by the dividend yield of $r - c$, which itself is the required rate of return minus the dividend growth rate.

We can rearrange (21.6) to give:

$$r = \frac{DV_1}{P_0} + c \tag{21.7}$$

which states that the share's fair value is the dividend yield together with the expected dividend growth rate. The expression at (21.7) is known as the *Gordon growth model* after its first presentation in Gordon (1962).

[3] This is not explained in detail here, but see for instance the appendix in chapter 5 of Higson (1995).

Expected earnings valuation model

In the expected earnings model, again the cash flow stream of the expected earnings is discounted at the required rate of return or market-determined discount rate *r*. To ensure consistency, the values given by the earnings model and the dividend model must be identical. This is achieved by using what are known as *economic earnings* rather than reported earnings when undertaking the valuation. Economic earnings of a share are defined as the maximum quantity of resources that may be withdrawn from the share and consumed before the share becomes unable to provide real consumption at a future date. A *cash flow statement* is used to convert reported earnings into economic earnings. Put simply, for sources of funds this is:

reported earnings + new external funds = total sources

while for uses of funds this is:

dividends + net investment = total uses

The relationship just given describes a *net* cash flow statement. If stated per share issued, the cash flow statement is given by

$$y_t + F_t = DV_t + x_t \tag{21.8}$$

where

y_t is the reported earnings per share in year t
F_t are the new external funds per share in year t
DV_t is the dividend per share in year t
x_t is the net investment per share in year t.

The relationship at (21.8) illustrates that if the firm raises new external finance, so that $F_t \neq 0$ a company can make the decision to proceed with new investment independently of the funding decision. However if all the new investment in the company is sourced from retained earnings, an increase in dividends will decrease net investment, which will then reduce the company's ability to generate further real income in the future.

Since they involve reinvesting a portion of its earnings and guaranteeing financial health, a company's retained earnings do not represent genuine economic income that is available to shareholders. From this reasoning, economic earnings per share are defined as $y_t + F_t - x_t$, where

$$\sum_{t=1}^{\infty} F_t / (1+r)^t = 0$$

We therefore reason that the present value of externally-sourced funds must be equal to zero, and as we assume this funding to be debt, all its debt is required to be repaid during the life of the company.

Therefore we say that:

economic earnings = reported earnings + new external funds − net investment

From this relationship and the expression above, we can value shares on the basis of economic earnings using the expression at (21.9):

$$P_0 = \sum_{t=1}^{\infty} \frac{E(y_t + F_t - x_t)}{(1+r)^t} = \sum_{t=1}^{\infty} \frac{E(y_t - x_t)}{(1+r)^t} \tag{21.9}$$

Identical valuations will be produced irrespective of the method used. This would be expected from no-arbitrage principles, but also because as (21.8) shows, the value of dividends would be equal to that of economic earnings.

Dividend policy

Some companies have never paid a dividend, a well-known example being Microsoft. However dividends analysis and dividend policy are fundamental tenets of corporate finance theory, so for this reason we consider it briefly here.

Classical finance writings such as Modigliani and Miller (1961) hold that a firm's dividend policy will not impact its valuation. The relationship at (21.8) would suggest that if a company can raise funds externally, its dividend policy will not affect its value. In theory and in practice, though, it can be shown to have some impact. Consider the dividend valuation model: under its approach a firm that did not pay dividends, generating return for its shareholders purely in the form of capital gain, would have a share value of zero. This is seen immediately by entering the DV_t value as 0. This is clear because under this model the entire value of the share would be contained at the infinity time point, which results in a present value of 0. So the conventional analysis states that the value of a share derives from its ability to cover dividend payments.[4] Even if no dividends are currently being paid, in order for shareholders to realise its value the firm will eventually have to pay dividends.

In practice the dividend policy adopted by a firm is analysed for its supposed information content. We summarise here the three conventional explanations why a firm might pay dividends.

Signalling policy

This explanation was first presented by Ross (1976). It suggests that dividend announcements are signals to the market of a company's intentions. As managers are privy to inside information on the true financial health of a company, it is logical for managers to signal this state of affairs by means of the dividend. So a

[4] As we noted, some well-known companies have never paid a dividend and have exhibited spectacular share price gains during their life; witness also the extraordinary bull run in 'dot.com' companies and the technology sector generally during the second half of the 1990s. However as there is as yet no formal model explaining this price behaviour, we confine ourselves to a review of the traditional approach.

company in good health might announce a dividend increase over last year's level. This would be a 'good' signal. However an increased dividend might also be a ruse by managers wishing to conceal the fact that their company was in parlous health. This is termed as a 'false' good signal.

The occurrence of false signals can be prevented by incentivising managers in an appropriate way, so that the penalty for sending false good signals is severe enough to lead instead to a true 'bad' signal being sent instead. A bad signal would be an unchanged dividend or a reduction in the dividend. For this be effective, signals also must be correlated with actual corporate performance, such that bad signals are associated with future insolvency. In this scenario a company that sent out false good signals would be faced with a greater chance of going bankrupt than if it had sent out a true bad signal. If these conditions exist, dividend levels and changes in these levels can be viewed as possessing significant information content on the state of the firm.

Principal–agent concern

This was presented by Rozeff (1977). Although shareholders are the owners of a firm, in practice they delegate the daily running of the firm to managers who run the company on the shareholders' behalf. Thus the shareholders are principals while the managers are agents. The interests of the two may conflict, however, leading to managers acting in ways that do not maximise shareholder value, but that are in their own short-term interest. To guard against this, the owners undertake monitoring, which has attendant agency costs. Such costs are minimised when funds are raised externally, as outside investors subject the firm to considerable inspection and scrutiny whenever this occurs. At the same time there are costs incurred when raising funds in such a way.

In this scenario, then, the payment of an annual dividend is viewed as a trade-off between agency costs and the costs of floatation. If dividends are at higher levels, externally sourced funds will have to be raised on a more frequent basis, and the average cost of so doing is raised as well. However the average agency costs decrease in this environment. So the 'optimum' dividend level for any company is that which minimises the sum of the two different costs.

Tax clientele effect

This was first presented by Modigliani and Miller (1961). A dividend policy might be used by a firm to dictate which class of investor ends up buying a company's shares. So investors liable to the top rate of tax will prefer to hold shares of low-dividend paying companies, while investors with low or zero marginal tax rates will prefer to hold the shares of high-dividend paying companies. Therefore if a company wishes particular classes of investor to hold its shares, it can tailor its dividend policy to effect this.

Under the supposed new paradigm, dividends are viewed as important to traditional 'bricks and mortar' companies, and considered unfashionable for, for example, technology companies. It is too early to state with certainty whether there has been such a paradigm shift, so we conclude for the moment that a dividend policy is important, and firms are likely to adopt one of the policies we have summarised here.

Selected bibliography and references

Choudhry, M. *The Bond and Money Markets: Strategy, Trading, Analysis*, Butterworth-Heinemann, 2001.

Gordon, M. *The Investment, Financing and Valuation of the Corporation*, Irwin, 1962.

Higson, C. *Business Finance*, 2nd edition, Butterworths, 1995.

Mills, R. and Robertson, J. *Fundamentals of Managerial Accounting and Finance*, Mars Business Associates, 1995.

Modigliani, F. and Miller, M. 'Dividend policy, growth and the valuation of shares', *Journal of Business*, 34, 1961, pp. 411–33.

Ross, S. 'The arbitrage theory of capital pricing', *Journal of Economic Theory*, December 1976, pp. 343–62.

Rozeff, M. 'The dividend discount model', *Journal of Portfolio Management*, Winter 1977.

Sharpe, W., Alexander, G. and Bailey, J. *Investments*, Prentice Hall, 1995.

Van Horne, J. *Financial Management and Policy*, 10th edn, Prentice Hall, 1995.

22
Introduction to Financial Ratio Analysis

The second chapter in our brief look at equity analysis considers some key concepts in finance, followed by an introduction to financial ratio analysis.

Key concepts in finance

The cornerstone of financial theory is the concept of the time value of money, which we introduced early in the book. This principle underpins discounted cash flow analysis, which is long established as a key element in financial analysis. The academic foundation of the *present value rule* as a corporate finance project appraisal technique is the Fisher–Hirshleifer model, generally quoted as first being presented by Fisher (1930) and Hirshleifer (1958). The present value rule established that in order to maximise shareholder wealth, a firm would be on safe ground accepting all projects with a positive net present value. The milestones in finance that followed this landmark are generally cited as being the *efficient markets hypothesis, portfolio theory* and the capital asset pricing model (CAPM). Without going into the mathematics and derivation, we briefly introduce these topics in this section.

The efficient market hypothesis is attributed to Fama (1965). Its primary message was that it is not possible to outperform the market. Investors who found they beat the market in the short term would not be able to sustain this over time because information reaches the market very quickly and other investors will react to this information immediately. This reaction of buying or selling assets will in turn impact share prices, so that shares rapidly become fully valued; after this only unexpected events will influence these prices. These events may have either a negative or positive impact on share prices, so that it becomes impossible to discern a clear trend in the movement of prices.

The key aspects of the efficient markets hypothesis are that:

- The current price of a stock reflects all that is known about the stock and the issuing company, as well as relevant market and economic information.
- If we accept market efficiency, share prices can be accepted to be fair value, and given all *publicly available* information, neither under or over-valued.

- It is not possible for an investor to beat the market, unless he is privy to information ahead of the market.
- Only unpredictable relevant news can cause share prices to change, and all previously released news has already been incorporated into the share price.
- Since unpredictable news is, by its nature, unpredictable, changes in share prices are also unpredictable and follow what is known as a *random walk*.

The key assumptions underlying the efficient market hypothesis are that investors are rational operators, and that being rational they will undertake dealing only on the receipt of new information, and not using intuition. The assumption of rationality later gave rise to the CAPM.

Portfolio theory was first presented by Markowitz (1959), and suggests that an investor who diversifies will achieve superior returns to one who does not. It also follows naturally from the efficient market hypothesis; as it is not possible to outperform the market, the most logical investment decision would be to hold the market itself in the form of a basket of shares that represent the entire market. The two main assumptions of the theory are that:

- Investment appraisal risk is given by the amount of variation in the returns over time.
- The overall risk level may be reduced if the assets are combined into a portfolio.

CAPM was developed from portfolio theory, and assumes that rational investors require a premium when holding risk-bearing assets. The model defines the risk premium of an individual share in relation to the market, and can be used as a project appraisal tool. The risk premium is measured by quantifying the volatility of an individual share in relation to the market as a whole, by means of the share's *beta*. Assuming that markets absorb all relevant information efficiently, share prices will react to information rapidly, and their adjusted price will then fully reflect all information received to date, as well as all expectations of the company's future prospects.

Individual shares are more or less risky than the average of the market as a whole, and this is captured by their beta. Regression techniques are commonly used to measure beta, using historical share price data. For example the London Business School's *Risk Measurement Service* uses monthly share price movements of the previous five years to estimate beta values for liquid securities.

The CAPM is attributed to Sharpe (1964), and states that:

- The return on a risk-bearing asset is the sum of the risk-free interest rate together with a risk premium, which is a multiple of the beta and the premium of the market itself.
- The constituents of share prices include their perceived risk-bearing level, and discounts built into them explain the higher returns achieved by certain investors.
- A portfolio of shares with high volatility will have a lower price for a specified return, so in order to generate higher returns, investors must accept higher risk.

An overview discussion of CAPM is given in Appendix 22.1.

Under CAPM the market itself has a beta value of 1. An individual share exhibiting identical price movements identical to the market therefore also has a beta of 1. A share that is three times as volatile as the market[1] will have a beta of 3.

A good practical example in the application and use of Beta is the South African gold market. Gold mining companies can have significantly different production costs. In the late 1980s and early 1990s a number of gold companies has production costs over US$400 per ounce. At current gold prices that would make those mines unprofitable, and some mines have in fact closed. All of the mines that operate today do so at a much reduced cost per ounce.

With a gold price hovering around $270 per ounce, a mining company (Company A) that is able to produce gold at $240 per ounce will make a profit of $30 per ounce. But a company (Company B) that is able to run at a lower cost of say $170 would in turn make a $100 profit per ounce. If the gold price were to increase by $20, that would increase Company A's profit by $20 or 67%. On the other hand, Company B's profit per ounce would only increase by 20%. This makes Company A much more sensitive or 'marginal' to movements in the gold price. In this case Company A would have a higher beta against gold than Company B.

This is discussed in a South African newspaper, the *Business Times*, in an article published in August 1998 entitled 'Gold fever hits Durban Roodepoort Deep at last'. The article quotes: 'Where a mine operates right at the margin and often at a loss, an uptick in gold leads to big jumps in profitability.'

The beta is the covariance of the securities return against the return of the market index (in our case gold), divided by the variance of the market index, or:

$$\beta = \frac{\mathrm{Cov}(r_s, r_m)}{\mathrm{Var}(r_m)} \tag{22.1}$$

In order to provide a practical working example, we calculated the beta of Roodepoort Deep over the period August 1998 to February 2001. As a comparison we also calculated the beta of Anglo Gold. The Anglo group of companies is the world's largest gold producer, and Anglo Gold Ltd has historically had a lower cost of production.

Although the calculation is simple, the method used to determine the return on the security or market could change the beta. In order to conceal any daily random movements we decided to use a 20-day moving average. It turns out that whether using the daily return or a 20-day moving average, the beta of Roodepoort Deep is approximately 1.5 times the beta of Anglo Gold.

This beta measure can be dependent on factors other than a mine's margin. Offshore mining will give rise to higher or lower currency exposure, and therefore a different return, and diversification in mining activities, for example into base and ferrous metals, would have an effect. To a certain extent, a company can even manage its own beta through effective treasury management. Companies that

[1]So that if the market rose by 10% the share price would rise by 30%, and a fall in the market of 10% would equally lead to a fall in the share price of 30%.

trade in gold forwards and options stand a much better chance of managing their beta in a way that suits management's and shareholders' objectives.

A critique of the CAPM and review of its strengths and weaknesses is outside the scope of this book. Comments about the effectiveness of beta as a measure of risk later led to the development of arbitrage pricing theory (APT), first presented by Ross (1976). This states that:

- Two assets possessing identical risk exposures must offer investors identical returns, otherwise an arbitrage opportunity will arise.
- The various elements of market risk can be measured in terms of a number of economic factors, including inflation levels, interest rates, production figures and so on, which influence all share prices.
- By using regression techniques it is possible to calculate an estimate of the impact of each of these economic factors on the overall level of risk.

There have been a number of criticisms of both the CAPM and APT,[2] and the valuation of companies that have yet to make a profit illustrates how analysts can no longer apply the traditional techniques to all companies. Essentially CAPM and APT assume that the past is a good representation of the future, which may be unrealistic for companies that have undergone or are undergoing significant changes, or that operate in rapidly changing or developing industries. As beta is measured by a regression of past returns over a relatively long period of time, any impact on the level of beta will be felt only slowly. Hence the historical beta of a company that has, say, changed its view on risk exposure, will not be a reliable estimate of its future beta. Finally, it is difficult to use either method for firms that have no publicly traded shares, or for divisions within companies. However no alternative models have been presented, which explains why both CAPM and APT and the efficient markets hypothesis are still generally used in the markets.

Ratio analysis

Ratio analysis is used heavily in financial analysis. In this section we present a review of the general application of ratio analysis and its use in peer group analysis.

Overview of ratio analysis

A number of performance measures are used as management information in the financial analysis of corporations. Generally they can be calculated from published accounts. The following key indicators are used by most listed companies to monitor their performance:

- return on capital employed
- profit on sales

[2] Most significantly, in Fama and French (1992). However see Roll and Ross (1992) for an argument that the CAPM and APT can still be applied.

- sales multiple on capital employed
- sales multiple of fixed assets
- sales per employee
- profit per employee.

These indicators are all related, and it is possible to measure the impact of an improvement in one of them on the others. Return on capital employed (ROCE) is defined in a number of ways, the two most common being return on net assets (RONA) and return on equity (ROE). RONA measures the overall return on capital irrespective of the long-term source of that capital, while ROE measures return on shareholders' funds only, thereby ignoring interest payments to providers of debt capital. To focus on RONA, which gives an indication of the return generated from net assets (that is, fixed assets and current assets minus current liabilities), analysts frequently split this into return on sales and sales multiples. Such measures are commonly calculated for quoted and unquoted companies, and are used in the comparison of performance between different companies.

We illustrate the calculation and use of these ratios in the next section.

Using ratio analysis

In Tables 22.1 and 22.2 we show the published accounts for a fictitious manufacturing company, Constructa plc (the notes to the accounts are in Table 22.3). These are the balance sheet and profit and loss account. From the information in the accounts we are able to calculate the RONA, return on sales and sales multiples ratios, shown in Table 22.4.

Table 22.1 Constructa plc balance sheet for the year ended 31 December 2000

	Notes	*2000* £m	*1999* £m	*1998* £m
Fixed assets		97.9	88.2	79.4
Current assets				
Stock		80.6	67.3	65.4
Debtors	(2)	44.3	40.5	39.6
Cash		2.4	2.7	1.4
		127.3	110.5	106.4
Creditors: amounts due within one year	(3)	104.8	85.8	70.0
Net current assets		22.5	24.7	36.4
Total assets less current liabilities		120.4	112.9	115.8
Creditors: amounts due after one year	(3)	31.4	36.9	35.5
		89.0	76.0	80.3
Capital and reserves				
Paid up share capital	(4)	15.0	15.0	15.0
Share premium account		45.5	37.2	46.1
Profit and loss account		28.5	23.8	19.2
Shareholders funds		89.0	76.0	80.3

Table 22.2 Constructa plc profit and loss account for the year ended 31 December 2000

	Notes	*2000* £m	*1999* £m	*1998* £m
Turnover		251.6	233.7	211.0
Cost of sales		118.2	109.3	88.7
Gross profit		133.4	124.4	122.3
Operating expenses		109.0	102.7	87.9
Operating profit		24.4	21.7	34.4
Interest payable	(1)	7.6	6.2	7.1
Profit before tax		16.8	15.5	27.3
Tax liability		5.04	4.65	8.19
Shareholders profit		11.8	10.9	19.1
Dividends		7.1	6.2	8.5
Reserves		4.7	4.6	10.6
Earnings per share		7.87	7.27	12.7

Table 22.3 Constructa plc: notes to the accounts

	2000 £m	*1999* £m	*1998* £m
(1) Interest payable			
Bank loans and short-term loans	5.8	4.1	5.4
Hire purchase	1.0	1.0	1.0
Leases and other loans	0.8	1.1	0.7
	7.6	6.2	7.1
(2) Debtors			
Trade debtors	34.3	31.8	32.1
Other debtors	10	8.7	7.5
	44.3	40.5	39.6
(3) Creditors: amounts due within one year			
Bank loans	31.7	26	21.1
Bond	7	7	7
Trade creditors	30.6	28.4	19.4
Tax and national insurance	10.8	6.8	3.8
Leases	3.5	2.6	11.7
Other creditors	8.9	4.1	1.4
Accruals	6.8	5.8	1.6
Dividend	5.5	5.1	4
	104.8	85.8	70.0
(3) Creditors: amounts due after one year			
Bank loans	12.1	11.8	10.2
Bond	7	7	7
Leases	8.9	9.4	9.1
Other creditors	3.4	8.7	9.2
	31.4	36.9	35.5
(4) Paid-up share capital			
10p ordinary shares, 150 million			

Table 22.4 Constructa plc RONA ratio measures

Ratio	*Calculation*	*2000*	*1999*	*1998*
RONA %	(3)/(5) × 100	20.3%	19.2%	29.7%
Return on sales %	(3)/(4) × 100	9.7%	9.3%	16.3%
Sales multiple (x)	(4)/(5)	2.1×	2.1×	1.8×
	Source	*£m*	*£m*	*£m*
(1) Profit before tax	P&L account	16.8	15.5	27.3
(2) Interest payable	P&L account	7.6	6.2	7.1
(3) Profit before interest and tax	(1) + (2)	24.4	21.7	34.4
(4) Sales ('turnover')	P&L account	251.6	233.7	211.0
(5) Net assets	Balance sheet	120.4	112.9	115.8

From Table 22.4 we see that Constructa's RONA measure was 20.3% in 2000; however on its own this figure is meaningless. In order to gauge the relative importance of this measure we would have to compare it with previous years' figures, to see if any trend was visible. Other useful comparisons would be to use the same measure for Constructa's competitor companies, as well as industry sector averages. From the information available here, it is possible only to make an historical comparison.

We see that the measure has fallen considerably from the 29.7% figure in 1998, but that the most recent year has improved from the year before. The sales margin shows exactly the same pattern; however, the sales generation figure has not decreased. During a period of falling return such as this, which is commonly encountered during a recession, a company would analyse its asset base, with a view to increasing the sales generation ratio and countering the decrease in decreasing margin ratio.

This illustration is a very basic one. Any management-level ratio analysis would need to look at a higher level if it is to provide any meaningful insight. We consider this in the next section.

Management-level ratio analysis

Return on equity

We now consider a number of performance measures that are used in corporate level analysis. Table 22.5 shows performance for a UK listed company in terms of return on equity (ROE). The terms we have considered, together with a few we have not, are shown as a historical trend. 'Asset turnover' refers to the sales generation or sales multiple, while 'leverage factor' is a measure of the *gearing* level, which we consider shortly.

Our analysis of the anonymous UK plc shows how ROE is linked to RONA, which we illustrated in the earlier analysis. How do the figures turn out for the hypothetical Constructa plc? These are listed in Table 22.6.

Table 22.5 A UK plc corporate performance 1995–9

Performance measure	*1999*	*1998*	*1997*	*1996*	*1995*
Asset turnover (sales generation)	2.01	1.97	1.85	1.91	1.79
Return on net sales	4.26%	4.43%	3.99%	4.77%	4.12%
Return on net assets*	8.56%	8.73%	7.38%	9.11%	7.37%
Leverage factor (gearing)	2.43	2.54	2.83	2.95	2.71
Return on equity~	20.80%	22.17%	20.89%	26.87%	19.97%

* This is Asset turnover × Return on net sales
~ This is Return on net assets × Leverage factor

Table 22.6 Constructa plc corporate-level ratios

Ratio	*Calculation*	*2000*	*1999*	*1998*
RONA %	See Table 22.1	20.3%	19.2%	29.7%
Return on sales %	See Table 22.1	9.7%	9.3%	16.3%
Sales multiple (×)	See Table 22.1	2.1×	2.1×	1.8×
ROE %	(6) / (7) × 100	13.26%	14.21%	23.78%
Gearing (×)	(5) / (7)	1.35×	1.49×	1.44×
	Source	*£m*	*£m*	*£m*
(1) Profit before tax	P&L account	16.8	15.5	27.3
(2) Interest payable	P&L account	7.6	6.2	7.1
(3) Profit before interest and tax	(1) + (2)	24.4	21.7	34.4
(4) Sales ('turnover')	P&L account	251.6	233.7	211.0
(5) Net assets	Balance sheet	120.4	112.9	115.8
(6) Shareholders' profit	P&L account	11.8	10.8	19.1
(7) Shareholders' funds	Balance sheet	89.0	76.0	80.3

Unlike our actual examples from the anonymous UK plc, the ratios for Constructa plc do not work out as a product of lower level ratios. This is because different profit measures have been used to calculate the RONA and ROE; it is deliberate. With RONA we wish to measure the profit generated by the business irrespective of the source of funds used in generating this profit. ROE on the other hand measures profit attributable to shareholders, so we use the 'profit after tax and interest' figure. The actual results illustrate a downtrend in the ROE, and senior management will be concerned about this. However this is outside the scope of this chapter. We consider gearing next.

Gearing

In Table 22.6 we encountered a leverage ratio, known as *gearing* in the UK. We also observed that gearing combined with RONA results in ROE. Put simply, gearing is the ratio of debt capital to equity capital, and measures the extent of indebtedness of a company. Gearing ratios are used by analysts and investors because they indicate the impact on ordinary shareholders' earnings of a change in operating profit.

For a company with high gearing, such a change in profit can have a disproportionate impact on shareholders' earnings because more of the profit has to be used to service debt. There is no one 'right' level of gearing, but at some point the level will be high enough to raise both shareholders' and rating agency concerns, as doubts creep in about the company's ability to meet its debt interest obligations.[3] The acceptable level of gearing for any company is dependent on a number of issues, including the type of business it is involved in, the average gearing level across similar companies, the stage of the business cycle (companies with high gearing levels are more at risk if the economy is heading into recession), the level of and outlook for interest rates, and so on. The common view is that a firm with an historically good track record that is less prone to the effects of changes in the business cycle can afford to be more highly geared than a company that does not boast these features.

As the values for debt and equity capital can be measured in more than one way, so a company's gearing level can take more than one value. We illustrate this below. Table 22.7 shows hypothetical company results.

From the data in Table 22.7 it is possible to calculate a number of different gearing ratios. These are shown in Table 22.8. So any individual measure of gearing is essentially meaningless unless it is also accompanied by a note of how it was calculated.

Market-book and price–earnings ratio

The remaining performance measures we wish to consider are the market-to-book ratio (MB) and the price–earnings or p/e ratio. It was not possible to calculate these for the hypothetical Constructa plc because we did not have a publicly quoted share price for it.[4] However these ratios are widely used and quoted by analysts and investors. For valuation purposes, they are used to obtain an estimated value of a

Table 22.7 Hypothetical company results

	£m
Short-term debt	190
Long-term debt	250
Preference shares	35
Shareholders' funds	500
Cash at bank	89
Market value of long-term debt	276
Market value of shareholders' funds	2,255

[3] A good illustration of this was the experience of telecommunications companies after they borrowed heavily in the debt capital market to pay for so-called 'third generation' mobile phone licences, which were auctioned off by different European governments. As a result of the multi-billion dollar sums involved in the purchase of each licence, some of the telecoms companies saw their credit ratings downgraded by Moodys and S&P (in the case of BT plc, to one level above non-investment grade) as concerns were raised about their resulting high gearing levels.

[4] Not every public listed company (plc) actually has its shares quoted on the stock exchange. It is possible for a company to be a plc without having quoted shares.

Table 22.8 Gearing ratios

Measure	*Gearing*
Long-term debt/Equity [250/(35 + 595)]	39.7%
Short-term and long-term debt/Shareholders' funds [(190 + 250) / 595]	74.0%
Short-term and long-term debt less cash at bank /Shareholders' funds [(190 + 250 − 89) / 595]	59.0%
Market value of long-term debt/Market value of equity [276 / 1977]	14.0%

company or subsidiary. Provided we have data for shareholders' earnings and shareholders' funds, as well as MB and p/e figures for comparable companies, it is possible to calculate an approximation of fair market value for an unquoted company.

The p/e ratio is considered to be an important performance indicator and for stock exchange listed companies is quoted in for example, the London *Financial Times*. It is given by

$$p/e = \frac{P_{share}}{EPS} \tag{22.2}$$

where P_{share} is the market price of the company's shares and *EPS* is the earnings per share. For quoted companies both these values may be obtained with ease.

The p/e ratio is an indication of the price that investors are prepared to pay for a company's shares in return for its current level of earnings. It relates shareholder profit to the market value of the company. Companies that are in 'high growth' sectors, such as (during the late 1990s) the 'dot.com' or technology sector, are observed generally to have high p/e ratios, while companies in low growth sectors will have lower p/e ratios. This illustrates one important factor in p/e ratio analysis: an individual figure on its own is of no real use. Rather, it is the sector average as well as the overall level of the stock market that are important considerations for the investor. In the *Financial Times* the company pages list the p/e ratio for each industry sector, thus enabling investors to compare specific company p/e ratios with the sector level and the market level.[5] The *p/e relative* is calculated by comparing specific and industry-level p/e ratios, given at (22.3), which is an indication of where investors rate the company in relation to the industry it is operating in, or the market as a whole.

$$p/e_{relative} = \frac{p/e_{company}}{p/e_{market}} \tag{22.3}$$

A very high p/e relative for a specific company may indicate a highly rated company and one that is a sector leader. However it may also indicate – and this is very

[5]These figures are not listed in the Monday edition of the *Financial Times*, which contains other relevant data.

topical – a 'glamour' company that is significantly overvalued and so overdue for a correction and decline in its share price.

The MB ratio relates a company's market value to shareholder funds value. If we see the p/e ratio as emanating from the P&L account, the MB ratio emanates from the balance sheet. It is given by

$$MB = ROE\ \% \times p/e\ ratio \tag{22.4}$$

We consider the MB and p/e ratios in the context of business valuation in the next section.

Corporate valuation

We have noted how for a company listed on a stock exchange, it is straightforward to know its market value: this is its share price. However for subsidiaries and divisions of quoted companies or unquoted companies, a proper market value is not so simple to obtain. In this section we provide an introduction of how analysis from within a 'peer group' of companies may be used to obtain an estimated valuation for unquoted companies.

We wish to calculate an estimated market share price for Constructa plc, our hypothetical manufacturing company. Assume that we are fortunate to observe a peer group that consists of three other manufacturing companies of comparable size and performance, operating in a similar line of business to Constructa plc. The three companies are known as X, Y and Z. Table 22.9 shows financial data and key performance indicators for the year 2000 for each of these three companies.

Table 22.9 Comparable company financial indicators, year 2000

	X plc	*Y plc*	*Z plx*
Turnover £m	821.4	369.7	211.3
Profit before interest and tax £m	97.6	41.9	18.7
Net profit (profit after interest and tax) £m	56.2	26.7	15.4
Book value of shareholders' funds £m	331.2	219.6	46.9
Shares in issue	167m	55m	48m
Share market price	712p	408p	926p
Return on sales %[1]	11.88	11.33%	8.85%
Earnings per share[2]	33.7p	48.5p	32.1p
p/e ratio[3]	21.1	8.4	28.8
Book value per share[4]	198p	399p	97.7p
MB ratio[5]	3.6	1.02	9.5

Notes
[1] Return on sales is profit before interest and tax/turnover
[2] EPS is net profit/number of shares in issue
[3] The p/e ratio is share price/earnings per share
[4] Book value per share is book value of shareholders' funds/number of shares in issue
[5] MB ratio is share market price/book value per share

Table 22.10 Peer group company ratios, mean values and Constructa plc market valuation

	Mean value	*X plc*	*Y plc*	*Z plx*
p/e ratio	19.4	21.1	8.4	28.8
MB ratio	4.7	3.6	1.02	9.5

Constructa plc		*Mean value*	*High value*	*Low value*
Valuation using p/e ratio				
EPS	7.87			
p/e ratio		19.4	28.8	8.4
Share market value[1]		152.7p	226.7p	66.1p
Valuation using MB ratio				
Book value per share	59.3p			
MB ratio		4.7	9.5	1.02
Share market value[2]		278.7p	563.4p	60.5p

Notes
[1] Obtained by multiplying EPS by p/e ratio
[2] Obtained by multiplying book value per share by MB ratio

The next step is to use this observed data in conjunction with Constructa plc data to obtain a range of possible values for the latter's market value. First, we calculate the mean p/e and MB ratios of the three peer group companies, and then from the range of ratios for these companies we calculate the estimated Constructa plc values, using that company's own earnings per share value. In this way, we obtain a highest and lowest possible market valuation and a mean valuation. We have not previously calculated a book value per share for Constructa plc, so this is done now; the result is 59.3 pence, obtained by dividing the shareholders' funds figure of £89 million by the number of shares (150 million).

The mean value p/e and MB ratios are shown in Table 22.10, together with the range of possible market values for Constructa plc using each method.

Through this approach we obtain a mean value for the Constructa plc share price of £1.53 or £2.79 depending on which method we use. It is a subjective issue which approach is the better one, and the motivation of the analyst undertaking the calculation is key. In practice, analysts will consider a peer group with a greater number of companies, which usually results in a wider range of possible values. Of course the true market valuation for any good is the price at which there is both a buyer and seller for it, and similarly the true value for a company will lie somewhere in between the high and low limits that arise from using the method we have just described.

Appendix 22.1: Capital asset pricing model

The capital asset pricing model (CAPM) is a cornerstone of modern financial theory and originates from analysis on the cost of capital. The cost of capital of a company may be broken down as shown by Figure 22.1.

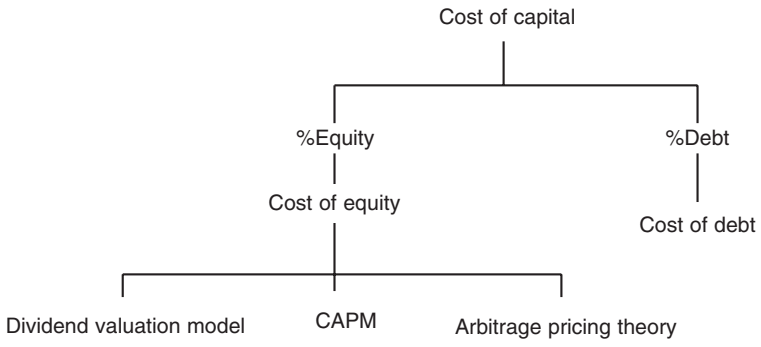


Figure 22.1 Components of the cost of capital

The three most common approaches used for estimating the cost of equity are the *dividend valuation model*, CAPM and the *arbitrage pricing theory*. CAPM is in a class of market models known as *risk premium* models, which rely on the assumption that every individual holding a risk-carrying security will demand a return in excess of the return he or she would receive from holding a risk-free security. This excess is the investor's compensation for her risk exposure. The risk premium in CAPM is measured by *beta*, and is known as *systematic, market* or *non-diversifiable* risk. This risk is caused by macroeconomic factors such as inflation or political events, which affect the returns of all companies. If a company is affected by these macroeconomic factors in the same way as the market (usually measured by a stock index), it will have a beta of 1, and will be expected to have returns equal to the market. Similarly if a company's systematic risk is greater than the market, then its capital will be priced such that it is expected to have returns greater than the market. Essentially therefore beta is a measure of volatility, with a company's relative volatility being measured by comparing its returns with the market's returns. For example if a share has a beta of 2.0, then on average for every 10% that the market index has returned above the risk-free rate, the share is expected to have returned 20%. Conversely for every 10% the market has under-performed the risk-free rate, the share is expected to have returned 20% below. Beta is calculated for a share by measuring its variance relative to the variance of a market index such as the FTSE All Share or the S&P 500. The most common method of estimating beta is with standard regression techniques based on historical share price movements over say, a five-year period.

To obtain the CAPM estimate of the cost of equity for a company, two other pieces of data are required, the risk-free interest rate and the equity risk premium. The risk-free rate represents the most secure return that can be achieved in the market. It is theoretically defined as an investment that has no variance and no covariance with the market. A perfect proxy for the risk-free rate therefore would be a security with a beta equal to zero, and no volatility. Such an instrument does

not, to all intents and purposes, exist. Instead the market uses the next-best proxy available, which in a developed economy is the government-issued Treasury bill, a short-dated debt instrument guaranteed by the government.

The equity risk premium represents the excess return above the risk-free rate that investors demand for holding risk-carrying securities. The risk premium in the CAPM is the premium above the risk-free rate on a portfolio assumed to have a beta of 1.0. The premium itself may be estimated in a number of ways. A common approach is to use historical prices, on the basis that past prices are a satisfactory guide to the future, and use these returns over time to calculate an arithmetic or geometric average. Research has shown that the market risk premium for the United States and UK has varied between 5.5% and 11% historically (Mills, 1994), depending on the time period chosen and the method used.

Once the beta has been determined, the cost of equity for a corporate is given by CAPM as (22.5):

$$k_e = r_f + (\beta \times r_e) \tag{22.5}$$

where

k_e is the cost of equity
r_f is the risk-free interest rate
r_e is the equity risk premium
β is the share beta.

The primary assumption behind CAPM is that all the market-related risk of a share can be captured in a single indicator, the beta. This would appear to be refuted by evidence that fund managers sometimes demand a higher return from one portfolio than another when both apparently are equally risky, with betas of 1.0. The difference in portfolio returns cannot be because of differences in specific risk, because diversification nearly eliminates such risk in large, well-balanced portfolios. If the systematic risk of the two portfolios were truly identical, then they would be priced to yield identical returns. Nevertheless the CAPM is often used by analysts to calculate cost of equity and hence cost of capital.

If we consider the returns on an individual share and the market as positively sloping lines on a graph plotting return, beta is usually given by (22.6):

$$r_s = \alpha_{sI} + \beta_{sI} r_I + \varepsilon_{sI} \tag{22.6}$$

where

r_s is the return on security s
r_I is the return on the market (usually measured for a given index)
α_{sI} is the intercept between s and I, often termed the 'alpha'
β_{sI} is the slope measurement or beta
ε is a random error term.

Selected references and bibliography

Fama, E. 'The behaviour of stock market prices', *Journal of Business*, January 1965, pp. 34–105.
Fama, E. and French, K. 'The cross-section of expected stock returns', *Journal of Finance*, June 1992, pp. 427–65.
Fisher, I. *Theory of Interest*, Macmillan, 1930.
Hirshleifer, J. 'On the theory of optimal investment decisions', *Journal of Political Economy*, 1958, pp. 329–72.
Markowitz, H. *Portfolio Selection*, Yale University Press, 1959.
Mills, R. *Strategic Value Analysis*, Mars Business Associates, 1994.
Roll, R. and Ross, S. A. 'On the cross-sectional relation between expected returns and betas', Yale School of Management Working Paper No. 21, 1992.
Ross, S. A. 'The arbitrage theory of capital pricing', *Journal of Economic Theory*, December 1976, pp. 343–62.
Sharpe, W. 'Capital asset prices: a theory of market equilibrium under conditions of risk', *Journal of Finance*, September 1964, pp. 425–42.

PART VI
Risk Measurement and Value-at-Risk

Financial market analysis must necessarily consider the risk exposure arising from participation in the markets. Investors and traders, by assuming either long or short positions in financial market instruments (whether exchange-traded or over-the-counter), will face a risk of loss arising from adverse movements in the markets. The most common risk measurement tool in use in the markets is known value-at-risk (VaR). Part VI of the book is an introduction to risk management for a bond or credit trader, and a look at the value-at-risk risk measurement tool.

This page intentionally left blank

23
Value-at-Risk and Credit VaR

In this chapter we review the main market risk measurement tool used in banking, known as value-at-risk (VaR). The review looks at the three main methodologies used to calculate VaR, as well as some of the key assumptions used in the calculations, including those on the normal distribution of returns, volatility levels and correlations. We also discuss the use of the VaR methodology with respect to credit risk.

Introducing value-at-risk

The introduction of VaR as an accepted methodology for quantifying market risk, and its adoption by bank regulators, are part of the evolution of risk management. The application of VaR has been extended from its initial use in securities houses to commercial banks and corporates, following its introduction in October 1994 when JP Morgan launched RiskMetrics™ free over the Internet.

VaR is a measure of the worst expected loss a firm might suffer over a period of time that has been specified by the user, under normal market conditions and a specified level of confidence. This measure may be obtained in a number of ways, using a statistical model or by computer simulation.

VaR is given over a specified time period for a set level of probability. For example if a daily VaR is stated as £100,000 to a 95% level of confidence, this means that during the day there is a only a 5% chance that the loss the next day will be *greater* than £100,000. VaR measures use estimated volatility and correlation. The 'correlation' referred to is the correlation that exists between the market prices of different instruments in a bank's portfolio. VaR is calculated within a given confidence interval, typically 95% or 99%; it seeks to measure the possible losses from a position or portfolio under 'normal' circumstances. The definition of normality is critical, and is essentially a statistical concept that varies by firm and by risk management system. Put simply, however, the most commonly used VaR models assume that the prices of assets in the financial markets follow a normal distribution. To implement VaR, all of a firm's positions data must be gathered into one centralised database. Once this is complete the overall risk has to be calculated by aggregating the risks from individual instruments across the entire portfolio.

The potential move in each instrument (that is, each risk factor) has to be inferred from past daily price movements over a given observation period. For regulatory purposes this period is at least one year. Hence the data on which VaR estimates are based should capture all relevant daily market moves over the previous year.

The main assumption underpinning VaR – and which in turn may be seen as its major weakness – is that the distribution of future price and rate changes will follow past variations. Therefore the potential portfolio loss calculations for VaR are worked out using distributions from historic price data in the observation period.

VaR is a measure of the volatility of a firm's banking or trading book. A portfolio containing assets that have a high level of volatility has a higher risk than one containing assets with a lower level of volatility. The VaR measure seeks to quantify in a single measure the potential losses that may be suffered by a portfolio.

VaR is therefore a measure of a bank's risk exposure; it is a tool for measuring market risk exposure. There is no one VaR number for a single portfolio, because different methodologies used for calculating VaR produce different results. The VaR number captures only those risks that can be measured in quantitative terms; it does not capture risk exposures such as operational risk, liquidity risk, regulatory risk or sovereign risk. It is important to be aware of what precisely VaR attempts to capture and what it clearly makes no attempt to capture. Also, VaR is not 'risk management'. A risk management department may choose to use a VaR measurement system in an effort to quantify a bank's risk exposure; however the application itself is merely a tool. Implementing such a tool in no way compensates for inadequate procedures and rules in the management of a trading book.

Assumption of normality

A distribution is described as *normal* if there is a high probability that any observation form the population sample will have a value that is close to the mean, and a low probability of a value far from the mean. The normal distribution curve is used by many VaR models, which assume that asset returns follow a normal pattern. A VaR model uses the normal curve to estimate the losses that an institution may suffer over a given time period. Normal distribution tables show the probability of a particular observation moving a certain distance from the mean.

If we look along a normal distribution table we see that at –1.645 standard deviations, the probability is 5%; this means that there is a 5% probability that an observation will be at least 1.645 standard deviations below the mean. This level is used in many VaR models.

Calculation methods

There are three different methods for calculating VaR. They are:

- the variance-covariance (or correlation or parametric method);
- historical simulation;
- Monte Carlo simulation.

Variance-covariance method

This method assumes the returns on risk factors are normally distributed, the correlations between risk factors are constant and the delta (or price sensitivity to changes in a risk factor) of each portfolio constituent is constant. Using the correlation method, the volatility of each risk factor is extracted from the historical observation period. Historical data on investment returns is therefore required. The potential effect of each component of the portfolio on the overall portfolio value is then worked out from the component's delta (with respect to a particular risk factor) and that risk factor's volatility.

There are several different methods of calculating the relevant risk factor volatilities and correlations. Two alternatives are:

- *Simple historic volatility*. This is the most straightforward method but the effects of a large one-off market move can significantly distort volatilities over the required forecasting period. For example if using 30-day historic volatility, a market shock will stay in the volatility figure for 30 days until it drops out of the sample range and correspondingly causes a sharp drop in (historic) volatility 30 days *after* the event. This is because each past observation is equally weighted in the volatility calculation.
- *To weight past observations unequally*. This is done to give more weight to recent observations, so that large jumps in volatility are not caused by events that occurred some time ago. One method is to use exponentially-weighted moving averages.

Historical simulation method

The historic simulation method for calculating VaR is the simplest, and avoids some of the pitfalls of the correlation method. Specifically the three main assumptions behind correlation (normally distributed returns, constant correlations, constant deltas) are not needed in this case. For historical simulation the model calculates potential losses using actual historical returns in the risk factors, and so captures the non-normal distribution of risk factor returns. This means rare events and crashes can be included in the results. As the risk factor returns used for revaluing the portfolio are actual past movements, the correlations in the calculation are also actual past correlations. They capture the dynamic nature of correlation as well as scenarios when the usual correlation relationships break down.

Monte Carlo simulation method

The third method, Monte Carlo simulation, is more flexible than the previous two. As with historical simulation, Monte Carlo simulation allows the risk manager to use actual historical distributions for risk factor returns rather than having to assume normal returns. A large number of randomly generated simulations are run forward in time using volatility and correlation estimates chosen by the risk manager. Each simulation will be different, but in total the simulations will aggregate to the chosen statistical parameters (that is, historical distributions and volatility

and correlation estimates). This method is more realistic than the previous two models, and therefore is more likely to estimate VaR more accurately. However its implementation requires powerful computers, and there is also a trade-off in that the time required to perform calculations is longer.

The level of confidence in the VaR estimation process is selected by the number of standard deviations of variance applied to the probability distribution. A standard deviation selection of 1.645 provides a 95% confidence level (in a one-tailed test) that the potential estimated price movement will not be more than a given amount based on the correlation of market factors to the position's price sensitivity.

Simple VaR calculation

To calculate the VaR for a single asset, we would calculate the standard deviation of its returns, using either its historical volatility or *implied volatility*. If a 95% confidence level is required, meaning we wish to have 5% of the observations in the left-hand tail of the normal distribution, this means that the observations in that area are 1.645 standard deviations away from the mean. This can be checked from standard normal tables. Consider the following statistical data for a government bond, calculated using one year's historical observations.

Nominal:	£10 million
Price:	£100
Average return:	7.35%
Standard deviation:	1.99%

The VaR at the 95% confidence level is 1.645 × 0.0199 or 0.032736. The portfolio has a market value of £10 million, so the VaR of the portfolio is 0.032736 × 10,000,000 or £327,360. So this figure is the maximum loss the portfolio may sustain over one year for 95% of the time.

We may extend this analysis to a two-stock portfolio. In a two-asset portfolio, we stated at (23.1) that there is a relationship that enables us to calculate the volatility of a two-asset portfolio; this expression is used to calculate the VaR, and is shown at (23.1):

$$Var_{port} = \sqrt{w_1^2 s_1^2 + w_{2s2}^2 {}^2 + 2 w_1 w_2 s_1 s_2 r_{1,2}} \tag{23.1}$$

where

w_1 is the weighting of the first asset
w_2 is the weighting of the second asset
s_1 is the standard deviation or *volatility* of the first asset
s_2 is the standard deviation or volatility of the second asset
$r_{1,2}$ is the correlation coefficient between the two assets.

In a two-asset portfolio the undiversified VaR is the weighted average of the individual standard deviations; the diversified VaR, which takes into account the

correlation between the assets, is the square root of the variance of the portfolio. In practice banks will calculate both diversified and undiversified VaR. The diversified VaR measure is used to set trading limits, while the larger undiversified VaR measure is used to gauge an idea of the bank's risk exposure in the event of a significant correction or market crash. This is because in a crash situation, liquidity dries up as market participants all attempt to sell off their assets. This means that the correlation relationship between assets ceases to have any impact on a book, as all assets move in the same direction. Under this scenario then, it is more logical to use an undiversified VaR measure.

Although the description given here is very simple, nevertheless it explains the essence of the VaR measure; VaR is essentially the calculation of the standard deviation of a portfolio, which is the used as an indicator of the volatility of that portfolio. A portfolio exhibiting high volatility will have a high VaR number. An observer might then conclude that the portfolio has a high probability of making losses. Risk managers and traders may use the VaR measure to help them to allocate capital to more efficient sectors of the bank, as return on capital can now be measured in terms of return on risk capital. Regulators may use the VaR number as a guide to the capital adequacy levels that they feel the bank requires.

Variance–covariance VaR

Calculation of variance–covariance VaR

In the previous section we showed how VaR could be calculated for a two-stock portfolio. Here we illustrate how this is done using matrices.

Consider the following hypothetical portfolio, invested in two assets, as shown in Table 23.1. The standard deviation of each asset has been calculated on historical observation of asset returns. Note that *returns* are returns of asset prices, rather than the prices themselves; they are calculated from the actual prices by taking the ratio of closing prices. The returns are then calculated as the logarithm of the price relatives. The mean and standard deviation of the returns are then calculated using standard statistical formulae. This would then give the standard deviation of

Table 23.1 Two-asset portfolio VaR

Assets	*Bond 1*	*Bond 2*	
Standard deviation	11.83%	17.65%	
Portfolio weighting	60%	40%	
Correlation coefficient			0.647
Portfolio value			£10,000,000
Variance			0.016506998
Standard deviation			12.848%
95% c.i. standard deviations			1.644853
VaR			0.211349136
VaR £			£2,113,491

daily price relatives, which is converted to an annual figure by multiplying it by the square root of the number of days in a year, usually taken to be 250.

The standard equation (shown as (23.1)) is used to calculate the variance of the portfolio, using the individual asset standard deviations and the asset weightings; the VaR of the book is the square root of the variance. Multiplying this figure by the current value of the portfolio gives us the portfolio VaR, which is £2,113,491.

The RiskMetrics VaR methodology uses matrices to obtain the same results as we have shown here. This is because once a portfolio starts to contain many assets, the method we described above becomes unwieldy. Matrices allow us to calculate VaR for a portfolio containing many hundreds of assets, which would require assessment of the volatility of each asset and correlations of each asset to all the others in the portfolio. We can demonstrate how the parametric methodology uses variance and correlation matrices to calculate the variance, and hence standard deviation, of a portfolio. The matrices are shown as Table 23.2. Note that the multiplication of matrices carries with it some unique rules; readers who are unfamiliar with matrices should refer to a standard mathematics textbook.

As shown in Table 23.2 using the same two-asset portfolio described, we can set a 2×2 matrix with the individual standard deviations inside; this is labelled the 'variance' matrix. The standard deviations are placed on the horizontal axis of the matrix, and a zero entered in the other cells. The second matrix is the correlation matrix, and the correlation of the two assets is placed in cells corresponding to the other asset; that is why a '1' is placed in the other cells, as an asset is said to have a correlation of 1 with itself. The two matrices are then multiplied to produce another matrix, labelled 'VC' in Table 23.2.[1]

Table 23.2 Matrix variance–covariance calculation for two-asset portfolio

	Variance matrix		*Correlation matrix*	
			Bond 1	Bond 2
Bond 1	11.83%	0	1	0.647
Bond 2	0	17.65%	0.647	1
	VC matrix		*Variance matrix*	
	0.1183	0.07654	11.83%	0
	0.1141955	0.1765	0	17.65%
	Weighting matrix		*VCV matrix*	
	60%	40%	0.013995	0.013509
			0.013509	0.031152
	WVCV		W	
	0.013801	0.020566	60%	
			40%	
			Standard deviation	

[1]Microsoft Excel has a function for multiplying matrices which may be used for any type of matrix. The function is '=MMULT()' typed in all the cells of the product matrix.

The VC matrix is then multiplied with the V matrix to obtain the variance-covariance matrix or VCV matrix. This shows the variance of each asset; for Bond 1 this is 0.01399, which is expected as that is the square of its standard deviation, which we were given at the start. The matrix also tells us that Bond 1 has a covariance of 0.0135 with Bond 2.

We then set up a matrix of the portfolio weighting of the two assets, and this is multiplied by the VCV matrix. This produces a 1×2 matrix, which we need to change to a single number, so this is multiplied by the W matrix, reset as a 2×1 matrix, which produces the portfolio variance. This is 0.016507. The standard deviation is the square root of the variance, and is 0.1284795 or 12.848%, which is what we obtained before.

In our illustration it is important to note the order in which the matrices were multiplied, as this will obviously affect the result. The volatility matrix contains the standard deviations along the diagonal, and zeros are entered in all the other cells. So if the portfolio we were calculating has 50 assets in it, we would require a 50×50 matrix and enter the standard deviations for each asset along the diagonal line. All the other cells would have a zero in them. Similarly for the weighting matrix, this is always one row, and all the weights are entered along the row. To take the example just given, the result would be a 1×50 weighting matrix.

The matrix method for calculating the standard deviation is more effective than the first method we described, because it can be used for a portfolio containing a large number of assets. In fact this is exactly the methodology used by RiskMetrics, and the computer model used for the calculation will be set up with matrices containing the data for hundreds, if not thousands, of different assets.

The variance–covariance method captures the diversification benefits of a multi-product portfolio because of the correlation coefficient matrix used in the calculation. For instance if the two bonds in our hypothetical portfolio had a negative correlation, the VaR number produced would be lower. It was also the first methodology introduced, by JP Morgan in 1994. To apply it, a bank would require data on volatility and correlation for the assets in its portfolio. This data is actually available from the RiskMetrics website (and other sources), so a bank does not necessarily need its own data. It might wish to use its own datasets however, should it have them, to tailor the application to its own use.

The advantages of the variance–covariance methodology are that it is simple to apply, and fairly straightforward to explain, and datasets for its use are immediately available. The drawbacks are that it assumes stable correlations and measures only linear risk; it also places excessive reliance on the normal distribution, and returns in the market are widely believed to have 'fatter tails' than a true normal distribution. This phenomenon is known as *leptokurtosis*, that is, the non-normal distribution of outcomes. Another disadvantage is that the process requires *mapping*. To construct a weighting portfolio for the RiskMetrics tool, cash flows from financial instruments are mapped into precise maturity points, known as grid points. We will review this later in the chapter. However in most cases assets do not fit into neat grid points, and complex instruments cannot be broken down accurately into cash flows. The mapping process makes assumptions that frequently do not hold in practice.

Nevertheless the variance–covariance method is still popular in the market, and is frequently the first VaR method installed at a bank.

Mapping

The cornerstone of variance–covariance methodologies such as RiskMetrics is the requirement for data on volatilities and correlations for assets in the portfolio. The RiskMetrics dataset does not contain volatilities for every maturity possible, as that would require a value for every period from 1 day to over 10,950 days (30 years) and longer, and correlations between each of these days. This would result in an excessive amount of calculation. Rather, volatilities are available for set maturity periods, and these are shown in Table 23.3.

If a bond is maturing in six years' time, its redemption cash flow will not match the data in the RiskMetrics dataset, so it must be mapped to two periods. In this case it would be split to the five-year and seven-year grid points. This is done in proportions so that the original value of the bond is maintained once it has been mapped. More importantly, when a cash flow is mapped, it must split in a manner that preserves the volatility characteristic of the original cash flow. Therefore, when mapping cash flows, if one cash flow is apportioned to two grid points, the share of the two new cash flows must equal the present value of the original cash flows, and the combined volatility of the two new assets must be equal to that of the original asset. A simple demonstration is given at Example 23.1.

Table 23.3 RiskMetrics grid points

1 month	3 months	6 months	1 year	2 years	3 years
4 years	5 years	7 years	9 years	10 years	15 years
20 years	30 years				

Example 23.1: Cash flow mapping

A bond trading book holds £1 million nominal of a gilt strip that is due to mature in precisely six years' time. To correctly capture the volatility of this position in the bank's RiskMetrics VaR estimate, the cash flow represented by this bond must be mapped to the grid points for five years and seven years, the closest maturity buckets for which the RiskMetrics dataset holds volatility and correlation data. The present value of the strip is calculated using the six-year zero-coupon rate, which RiskMetrics obtains by interpolating between the five-year rate and the seven-year rate. The details are shown in Table 23.4.

Note that the correlation between the two interest rates is very close to 1; this is expected because five-year interest rates generally move very closely in line with seven-year rates.

Table 23.4 Bond position to be mapped to grid points

Gilt strip nominal (£)	1,000,000
Maturity (years)	6
5-year zero-coupon rate	5.35%
7-year zero-coupon rate	5.50%
5-year volatility	24.50%
7-year volatility	28.95%
Correlation coefficient	0.979

We wish to assign the single cash flow to the five-year and seven-year grid points (also referred to as *vertices*). The present value of the bond, using the six-year interpolated yield, is £728,347. This is shown in Table 23.5, which also uses an interpolated volatility to calculate the volatility of the six-year cash flow. However we wish to calculate a portfolio volatility based on the apportionment of the cash flow to the five-year and seven-year grid points. To do this, we need to use a weighting to allocate the cash flow between the two vertices. In the hypothetical situation used here, this presents no problem because six years falls precisely between five years and seven years. Therefore the weightings are 0.5 for year five and 0.5 for year seven. If the cash flow had fallen in less obvious a maturity point, we would have to calculate the weightings using the formula for portfolio variance. Using these weightings, we calculate the variance for the new 'portfolio', containing the two new cash flows, and then the standard deviation for the portfolio. This gives us a VaR for the strip of £265,853.

Table 23.5 Cash flow mapping and portfolio variance

Interpolated yield	0.05425
Interpolated volatility	0.26725
Present value	728,347.0103
Weighting 5-year grid point	0.5
Weighting 7-year grid point	0.5
Variance of portfolio	0.070677824
Standard deviation	0.265853012
VaR £	265,853

Confidence intervals

Many models estimate VaR at a given confidence interval, under normal market conditions. This assumes that market returns generally follow a random pattern, but one that approximates over time to a normal distribution. The level of confidence at which the VaR is calculated will depend on the nature of the trading book's activity and what the VaR number is being used for. The Market Risk

amendment to the Basle Capital Accord stipulates a 99% confidence interval and a ten-day holding period if the VaR measure is to be used to calculate the regulatory capital requirement. However certain banks prefer to use other confidence levels and holding periods; the decision on which level to use is a function of asset types in the portfolio, quality of market data available, and the accuracy of the model itself, which will have been tested over time by the bank.

For example, a bank may view a 99% confidence interval as providing no useful information, as it implies that there should only be two or three breaches of the VaR measure over the course of one year. That would leave no opportunity to test the accuracy of the model until a relatively long period of time had elapsed, and in the meantime the bank would be unaware if the model was generating inaccurate numbers. A 95% confidence level implies the VaR level being exceeded around one day each month, if a year is assumed to contain 250 days.[2]

If a VaR calculation is made using 95% confidence, and a 99% confidence level is required for say regulatory purposes, we need to adjust the measure to take account of the change in standard deviations required. For example, a 99% confidence interval corresponds to 2.32 standard deviations, while a 95% level is equivalent to 1.645 standard deviations. Thus to convert from 95% confidence to 99% confidence, the VaR figure is divided by 1.645 and multiplied by 2.32.

In the same way there may be occasions when a firm will wish to calculate VaR over a different holding period from that recommended by the Basle Committee. The holding period of a portfolio's VaR calculation should represent the period of time required to unwind the portfolio, that is, sell off the assets on the book. A ten-day holding period is recommended but would be unnecessary for a highly liquid portfolio, for example one holding government bonds.

To adjust the VaR number to fit it to a new holding period we simply scale it upwards or downward by the square root of the time period required. For example a VaR calculation measured for a ten-day holding period will be $\sqrt{10}$ times larger than the corresponding one-day measure.

Historical VaR methodology

The historical approach to VaR is a relatively simple calculation, and it is also easy to implement and explain. To implement it, a bank requires a database record of its past profit/loss figures for the total portfolio; the required confidence interval is then applied to this record, to obtain a cut-off of the worst-case scenario. For example, to calculate the VaR at a 95% confidence level, the fifth percentile in value for the historical data is taken, and this is the VaR number. For a 99% confidence level measure, the 1% percentile is taken. The advantage of the historical method is that it uses the actual market data that a bank has recorded (unlike RiskMetrics, for example, for which the volatility and correlations are not actual

[2] For the 99% confidence level, $250 \times 1\% = 2.5$ days in one year, while 95% confidence is $250 \times 5\%$ or 12.5 days.

values, but estimated values calculated from average figures over a period of time, usually the last five years), and so produces a reasonably accurate figure. Its main weakness is that as it is reliant on actual historical data built up over a period of time; generally at least one year's data is required to make the calculation meaningful. Therefore it is not suitable for portfolios whose asset weightings frequently change, as another set of data would be necessary before a VaR number could be calculated.

In order to overcome this drawback banks use a method known as historical simulation. This calculates VaR for the current portfolio weighting, using the historical data for the securities in the current portfolio. To calculate historical simulation VaR for our hypothetical portfolio considered earlier, comprising 60% of bond 1 and 40% of bond 2, we require the closing prices for both assets over the specified previous period (usually three or five years); we then calculate the value of the portfolio for each day in the period assuming constant weightings.

Simulation methodology

The most complex calculations use computer simulations to estimate value-at-risk. The most common is the Monte Carlo method. To calculate VaR using a Monte Carlo approach, a computer simulation is run in order to generate a number of random scenarios, which are then used to estimate the portfolio VaR. The method is probably the most realistic, if we accept that market returns follow a similar 'random walk' pattern. However Monte Carlo simulation is best suited to trading books containing large option portfolios, whose price behaviour is not captured very well with the RiskMetrics methodology. The main disadvantage of the simulation methodology is that it is time-consuming and uses a substantial amount of computer resources.

A Monte Carlo simulation generates simulated future prices, and it may be used to value an option as well as for VaR applications. When used for valuation, a range of possible asset prices are generated and these are used to assess what intrinsic value the option will have at those asset prices. The present value of the option is then calculated from these possible intrinsic values. Generating simulated prices, although designed to mimic a 'random walk', cannot be completely random because asset prices, although not a pure normal distribution, are not completely random either. The simulation model is usually set to generate very few extreme prices. Strictly speaking, it is asset price returns that follow a normal distribution, or rather a lognormal distribution.

Monte Carlo simulation may also be used to simulate other scenarios, for example the effect on option 'greeks' for a given change in volatility, or any other parameters. The scenario concept may be applied to calculating VaR as well. For example, if 50,000 simulations of an option price are generated, the 95th lowest value in the simulation will be the VaR at the 95% confidence level. The correlation between assets is accounted for by altering the random selection programme to reflect relationships.

Example 23.2: Portfolio volatility using variance-covariance and simulation methods

A simple two-asset portfolio is composed of the following instruments:

	Gilt strip	FTSE100 stock
Number of units	£100 million	5 million
Market value	£54.39 million	£54 million
Daily volatility	£0.18 million	£0.24 million

The correlation between the two assets is 20%. Using (23.4) we calculate the portfolio VaR as follows:

$$Vol = \sqrt{s^2_{\text{bond}} + s^2_{\text{stock}} + 2 s_{bond} s_{stock} r_{bond,stack}}$$

$$Vol = \sqrt{0.18^2 + 0.24^2 + (2 \times 0.18 \times 0.24 \times 0.2)} = 0.327$$

We have ignored the weighting element for each asset because the market values are roughly equal. The calculation gives a portfolio volatility of £0.327 million. For a 95% confidence level VaR measure, which corresponds to 1.645 standard deviations (in a one-tailed test) we multiply the portfolio volatility by 1.645, which gives us a portfolio value-at-risk of £0.538 million.

In a Monte Carlo simulation we also calculate the correlation and volatilities of the portfolio. These values are used as parameters in a random number simulation to throw out changes in the underlying portfolio value.

Table 23.6 Monte Carlo simulation results

Simulation	*Market value: bond*	*Market value: stock*	*Portfolio value*	*Profit/loss*
1	54.35	54.9	109.25	0.86
2	54.64	54.02	108.66	0.27
3	54.4	53.86	108.26	−0.13
4	54.25	54.15	108.4	0.01
5	54.4	54.17	108.57	0.18
6	54.4	54.03	108.43	0.04
7	54.31	53.84	108.15	−0.24
8	54.3	53.96	108.26	−0.13
9	54.46	54.11	108.57	0.18
10	54.32	53.92	108.24	−0.15
11	54.31	53.97	108.28	−0.11
12	54.47	54.08	108.55	0.16
13	54.38	54.03	108.41	0.02
14	54.71	53.89	108.6	0.21
15	54.29	54.05	108.34	−0.05

These values are used to reprice the portfolio, and this value will be either a gain or loss on the actual mark-to-market value. This process is repeated for each random number that is generated.

In Table 23.6 we show the results for 15 simulations of our two-asset portfolio. From the results we read off the loss level that corresponds to the required confidence interval.

As the number of trials is increased, the results from a Monte Carlo simulation approach those of the variance-covariance measure. This is shown in Figure 23.1.

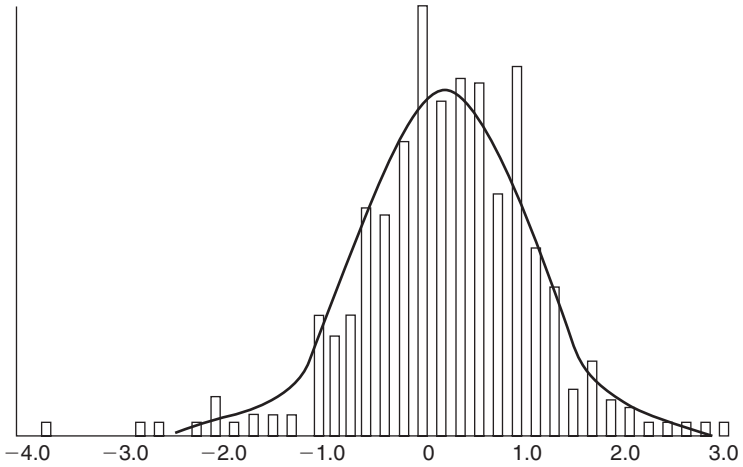


Figure 23.1 The normal approximation of returns

VaR for fixed income instruments

Perhaps the most straightforward instruments to which VaR can be applied are foreign exchange and interest-rate instruments such as money-market products, bonds, forward-rate agreements and swaps. In this section we review the calculation of VaR for a sample portfolio of bonds.

Sample bond portfolio

Table 23.7 details the bonds that are in our portfolio; for simplicity we assume that all the bonds pay an annual coupon and have full years left to maturity. In order to calculate the value-at-risk we first need to value the bond portfolio itself. The bonds are valued by breaking them down into their constituent cash flows; the present value of each cash flow is then calculated, using the appropriate zero-coupon interest rate. Note from Figure 23.2 that the term structure is inverted.

Table 23.7 Sample three-bond portfolio

	Bond 1	*Bond 2*	*Bond 3*
Nominal value	10,000,000	3,800,000	9,700,000
Coupon	5%	7.25%	6%
Maturity	5	7	2

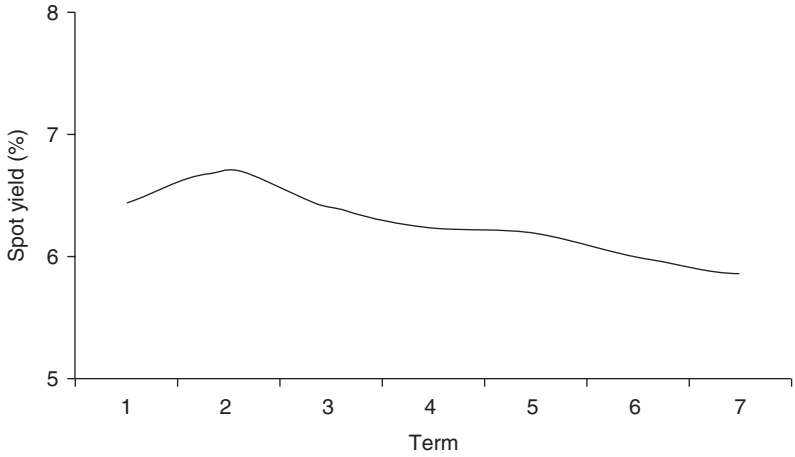


Figure 23.2 Term structure used in the valuation

Table 23.8 shows the present values for each of the cash flows. The total portfolio value is also shown.

We then use the volatility for each period rate to calculate the VaR. Data on interest-rate volatility is available for example, from the RiskMetrics website for all major currencies. The volatility levels for our hypothetical currency are relatively low in this example. The VaR for each maturity period is then obtained by multiplying the total present value of the cash flows for that period by its volatility level. This is shown in Table 23.9. By adding together all the individual values, we obtain an undiversified VaR for the portfolio. The total VaR is £1.77 million, for a portfolio with a market value of £23.1 million.

The figure just calculated is the undiversified VaR for the bond portfolio. To obtain the diversified VaR for the book, we require the correlation coefficient of each interest rate with the other interest rates (the correlation will be very close to unity, although the shorter-dated rates will be closer in line with each other than with long-dated rates). We may then use the standard variance-covariance approach, using a matrix of the undiversified VaR values and a matrix with the correlation values. However the diversification benefit of a portfolio of bonds will be small, mainly because their volatilities will be closely correlated.

Table 23.8 Bond portfolio valuation

Period	*Cash flows Bond 1*	*Bond 2*	*Bond 3*	*Zero-coupon rates*	*Discount factor*	*Present values Bond 1*	*Bond 2*	*Bond 3*
1	500,000	275,500	582,000	6.45	0.939408173	469,704	258,807	546,736
2	500,000	275,500	10,282,000	6.7	0.878357191	439,179	241,987	9,031,269
3	500,000	275,500		6.4	0.830185447	415,093	228,716	
4	500,000	275,500		6.25	0.784664935	392,332	216,175	
5	10,500,000	275,500		6.18	0.740945722	7,779,930	204,131	
6		275,500		5.98	0.705759136	0	194,437	
7		4,075,500		5.87	0.670794678	0	2,733,824	
					Totals	9,496,238	4,078,077	9,578,004
					Portfolio value	23,152,319		

Table 23.9 Bond portfolio undiversified VaR

Period	*Cash flows*	*Present value*	*Volatility*	*VaR*
1	1,357,500	1,275,246.595	0.0687	87,609.44105
2	11,057,500	9,712,434.639	0.0695	675,014.2074
3	775,500	643,808.8143	0.07128	45,890.69228
4	775,500	608,507.6568	0.0705	42,899.7898
5	10,775,500	7,984,060.629	0.08501	678,724.9941
6	275,500	194,436.642	0.08345	16,225.73778
7	4,075,500	2,733,823.711	0.08129	222,232.5295
Undiversified VaR				1768597.392

Forward-rate agreements

The VaR calculation for a *forward-rate agreement* (FRA) follows the same principles reviewed in the previous section. An FRA is a notional loan or deposit for a period starting at some point in the future; in effect it is used to fix a borrowing or lending rate. The derivation of an FRA rate is based on the principle of what it would cost for a bank that traded one to hedge it; this is known as the 'breakeven' rate. So a bank that has bought a 3v6 FRA (referred to as a 'threes-sixes FRA') has effectively borrowed funds for three months and placed the funds on deposit for six months. Therefore a FRA is best viewed as a combination of an asset and a liability, and that is how one is valued. So a long position in a 3v6 FRA is valued as the present value of a three-month cash-flow asset and the present value of a six-month cash-flow liability, using the three-month and six-month deposit rates. The net present value is taken, of course, because one cash flow is an asset and the other a liability.

Consider a 3v6 FRA that has been dealt at 5.797%, the three-month forward-forward rate. The value of its constituent (notional) cash flows is shown in Table 23.10 (overleaf). The three-month and six-month rates are cash rates in the market, while the interest-rate volatilities have been obtained from RiskMetrics. The details are summarised in Table 23.10.

The undiversified VaR is the sum of the individual VaR values, and is £34,537. It has little value in the case of a FRA, however, and would overstate the true VaR, because an FRA is made up of a notional asset and liability, so a fall in the value of one would see a rise in the value of the other. Unless a practitioner was expecting three-month rates to go in an opposite direction to six-month rates, there is

Table 23.10 Undiversified VaR for 3 v 6 FRA

Cash flow	*Term (days)*	*Cash rate*	*Interest rate*	*Present value volatilities*	*Undiversified VaR*
10,000,000	91	5.38%	0.14%	9,867,765	13,815
−10,144,536	182	5.63%	0.21%	−9,867,765	20,722

an element of diversification benefit. There is a high correlation between the two rates, so the more logical approach is to calculate a diversified VaR measure.

For an instrument such as an FRA, the fact that the two rates used in calculating the FRA rate are closely positively correlated will mean that the diversification effect will be to reduce the VaR estimate, because the FRA is composed notionally of an asset and a liability. From the values in Table 23.10 therefore, the six-month VaR is actually a negative value (if the bank had sold the FRA, the three-month VaR would have the negative value). To calculate the diversified VaR then requires the correlation between the two interest rates, which may be obtained from the RiskMetrics dataset. This is observed to be 0.87. This value is entered into a 2×2 correlation matrix and used to calculate the diversified VaR in the normal way. The procedure is:

- Transpose the weighting VaR matrix, to turn it into a 2×1 matrix.
- Multiply this by the correlation matrix.
- Multiply the result by the original 1×2 weighting matrix.
- This gives us the variance; the VaR is the square root of this value.

The result is an diversified VaR of £11,051.

Interest-rate swaps

To calculate a variance–covariance VaR for an interest-rate swap, we use the process described earlier for an FRA. There are more cash flows that go to make up the undiversified VaR, because a swap is essentially a strip of FRAs. In a plain vanilla interest-rate swap, one party pays fixed rate basis on an annual or semi-annual basis, and receives floating-rate interest, while the other party pays floating-rate interest payments and receives fixed-rate interest. Interest payments are calculated on a notional sum, which does not change hands, and only interest payments are exchanged. In practice, it is the net difference between the two payments that is transferred.

The fixed rate on an interest-rate swap is the breakeven rate that equates the present value of the fixed-rate payments to the present value of the floating-rate payments; as the floating-rate payments are linked to a reference rate such as Libor, we do not know what they will be, but we use the forward rate applicable to each future floating payment date to calculate what it would be if we were to fix it today. The forward rate is calculated from zero-coupon rates today. A 'long' position in a swap is to pay fixed and receive floating, and is conceptually the same as being short in a fixed-coupon bond and being long in a floating-rate bond; in effect the long is 'borrowing' money, so a rise in the fixed rate will result in a rise in the value of the swap. A 'short' position is receiving fixed and paying floating, so a rise in interest rates results in a fall in the value of the swap. This is conceptually similar to a long position in a fixed-rate bond and a short position in a floating-rate bond.

Describing an interest-rate swap in conceptual terms of fixed and floating-rate bonds gives some idea as to how it is treated for value-at-risk purposes. The coupon

Table 23.11 Fixed-rate leg of five-year interest rate swap and undiversified VaR

Pay date	*Swap rate*	*Principal (£)*	*Coupon (£)*	*Coupon present value (£)*	*Volatility*	*Undiversified VaR*
07 June 00	6.73%	10,000,000	337,421	327,564	0.05%	164
07 Dec 00	6.73%	10,000,000	337,421	315,452	0.05%	158
07 June 01	6.73%	10,000,000	335,578	303,251	0.10%	303
07 Dec 01	6.73%	10,000,000	337,421	294,898	0.11%	324
07 June 02	6.73%	10,000,000	335,578	283,143	0.20%	566
09 Dec 02	6.73%	10,000,000	341,109	277,783	0.35%	972
09 June 03	6.73%	10,000,000	335,578	264,360	0.33%	872
08 Dec 03	6.73%	10,000,000	335,578	256,043	0.45%	1,152
07 June 04	6.73%	10,000,000	335,578	248,155	0.57%	1,414
07 Dec 04	6.73%	10,000,000	337,421	242,161	1.90%	4,601
Total						10,528

on a floating-rate bond is reset periodically in line with the stated reference rate, usually Libor. Therefore the duration of a floating-rate bond is very low, and conceptually the bond may be viewed as being the equivalent of a bank deposit, which receives interest payable at a variable rate. For market risk purposes,[3] the risk exposure of a bank deposit is nil, because its present value is not affected by changes in market interest rates. Similarly, the risk exposure of a floating-rate bond is very low and to all intents and purposes its VaR may be regarded as zero. This leaves only the fixed-rate leg of a swap to measure for VaR purposes.

Table 23.11 shows the fixed-rate leg of a five-year interest-rate swap. To calculate the undiversified VaR we use the volatility rate for each term interest rate; this may be obtained from RiskMetrics. Note that the RiskMetrics dataset supports only liquid currencies: for example, data on volatility and correlation is not available for certain emerging market economies. We show the VaR for each payment; the sum of all the payments constitutes the undiversified VaR. We then require the correlation matrix for the interest rates, and this is used to calculate the diversified VaR. The weighting matrix contains the individual term VaR values, which must be transposed before being multiplied by the correlation matrix.

Using the volatilities and correlations supplied by RiskMetrics the diversified VaR is shown to be £10,325. This is very close to the undiversified VaR of £10,528. This is not unexpected because the different interest rates are very closely correlated.

Using VaR to measure market risk exposure for interest rate products enables a risk manager to capture non-parallel shifts in the yield curve, which is as

[3] We emphasise for *market* risk purposes; the credit risk exposure for a floating-rate bond position is a function of the credit quality of the issuer.

advantage over the traditional duration and interest-rate gap measures. Therefore estimating a book's VaR measure is useful not only for the trader and risk manager, but also for senior management, who by using VaR will have a more accurate idea of the risk market exposure of the bank. VaR methodology captures pivotal shifts in the yield curve by using the correlations between different maturity interest rates; this reflects the fact that short-term interest rates and long-term interest rates are not perfectly positively correlated.

Stress testing

Risk measurement models and their associated assumptions are not without limitation. It is important to understand what will happen should some of the model's underlying assumptions break down. *Stress testing* is a process whereby a series of scenario analyses or simulations are carried out to investigate the effect of extreme market conditions on the VaR estimates calculated by a model. It is also an analysis of the effect of violating any of the basic assumptions behind a risk model. If carried out efficiently, stress testing will provide clearer information on the potential exposures at risk due to significant market corrections, which is why the Basel Committee recommends that it be carried out.

Simulating stress

There is no standard way to undertake stress testing. It is a means of experimenting with the limits of a model; it is also a means to measure the residual risk which is not effectively captured by the formal risk model, thus complementing the VaR framework. If a bank uses a confidence interval of 99% when calculating its VaR, the losses on its trading portfolio due to market movements should not exceed the VaR number on more than one day in 100. For a 95% confidence level the corresponding frequency is one day in 20, or roughly one trading day each month. The question to ask is, 'What are the expected losses on those days?' Also, what can an institution do to protect itself against these losses?

The assumption that returns are normally distributed provides a workable daily approximation for estimating risk, but when market moves are more extreme, these assumptions no longer add value. The 1% of market moves that are not used for VaR calculations contain events such as the October 1987 crash, the bond market collapse of February 1994 and the Mexican peso crisis at the end of 1994. In these cases market moves were much larger than any VaR model could account for; in fact the October 1987 crash was a 20 standard deviation move. Under these circumstances correlations between markets also increase well above levels normally assumed in models.

An approach used by risk managers is to simulate extreme market moves over a range of different scenarios. One method is to use Monte Carlo simulation. This allows dealers to push the risk factors to greater limits; for example a 99% confidence interval captures events up to 2.33 standard deviations from the mean asset return level. A risk manager can calculate the effect on the trading portfolio of a

10 standard deviation move. Similarly risk managers may want to change the correlation assumptions under which they normally work. For instance if markets all move down together, something that happened in Asian markets from the end of 1997 and emerging markets generally from July 1998 after the Russian bond technical default, losses will be greater than if some markets are offset by other negatively correlated markets.

Only by pushing the bounds of the range of market moves that are covered in the stress testing process can financial institutions have an improved chance of identifying where losses might occur, and therefore a better chance of managing their risk effectively.

Stress testing in practice

For effective stress testing, a bank has to consider non-standard situations. The following are examples:

- parallel yield curve shifts of 100 basis points up and down
- steepening and flattening of the yield curve (2-year to 10-year) by 25 basis points
- increase and decrease in 3-month yield volatilities by 20%
- increase and decrease in equity index values by 10%
- increase and decrease in swap spread by 20 basis point.

These scenarios represent a starting point for a framework for routine stress testing.

Banks agree that stress testing must be used to supplement VaR models. The main problem appears to be difficulty in designing appropriate tests. The main issues are:

- difficulty in 'anticipating the unanticipated';
- adopting a systematic approach, with stress testing carried out by looking at past extremes and analysing the effect on the VaR number under these circumstances;
- selecting 10 scenarios based on past extreme events and generating portfolio VaRs based on re-runs of these scenarios.

The latest practice is to adapt stress tests to suit the particular operations of a bank. On the basis that one of the main purposes of stress testing is to provide senior management with accurate information of the extent of a bank's potential risk exposure, more valuable data will be gained if the stress test is particularly relevant to the bank. For example, an institution such as Standard Chartered Bank, which has a relatively high level of exposure to exotic currencies, may design stress tests that take into account extreme movements in, say, regional Asian currencies. A mortgage book holding option positions only to hedge its cash book, say one of the former UK building societies that subsequently converted to banks, may have no need for excessive stress testing on (for example) the effect of extreme moves in derivatives liquidity levels.

VAR methodology for credit risk

Credit risk emerged as a significant risk management issue during the 1990s and again during the 2007–8 credit crunch. In increasingly competitive markets, banks and securities houses began taking on greater credit risk in this period.

The growth in credit exposures and rise of complex instruments led to a need for more sophisticated risk management techniques.

Credit risk

There are two main types of credit risk, credit spread risk and credit default risk.

Credit spread risk

Credit spread is the excess premium, over and above government or risk-free risk, required by the market for taking on a certain assumed credit exposure. Credit spread risk is the risk of financial loss resulting from changes in the level of credit spreads used in the marking-to-market of a product. It is exhibited by a portfolio for which the credit spread is traded and marked. Changes in observed credit spreads affect the value of the portfolio and can lead to losses for investors.

Credit default risk

This is the risk that an issuer of debt (obligor) is unable to meet its financial obligations. Where an obligor defaults, a firm generally incurs a loss equal to the amount owed by the obligor less any recovery amount which the firm recovers as a result of foreclosure, liquidation or restructuring of the defaulted obligor. All portfolios of exposures exhibit credit default risk.

The VaR measurement methodology was first applied for credit risk by JPMorgan, which introduced the *CreditMetrics* tool in 1995. The measurement of credit risk requires a slightly different approach from that used for market risk, because the distribution of credit losses follows a different pattern from market risk. In the following sections we describe the approach used to measure such risk.

Modelling credit risk

Credit risk VaR methodologies take a portfolio approach to credit risk analysis. This means that:

- credit risks to each obligor across the portfolio are re-stated on an equivalent basis and aggregated in order to be treated consistently, regardless of the underlying asset class
- correlations of credit quality moves across obligors are taken into account.

This allows portfolio effects – the benefits of diversification and risks of concentration – to be quantified.

The portfolio risk of an exposure is determined by four factors:

- size of the exposure
- maturity of the exposure
- probability of default of the obligor
- systematic or concentration risk of the obligor.

Credit VaR, like market risk VaR, considers (credit) risk in a mark-to-market framework. It arises from changes in value due to credit events, that is, changes in obligor credit quality including defaults, upgrades and downgrades.

Nevertheless credit risk is different in nature from market risk. Typically market return distributions are assumed to be relatively symmetrical and approximated by normal distributions. In credit portfolios, value changes will be relatively small upon minor up/down grades, but can be substantial upon default. This remote probability of large losses produces skewed distributions with heavy downside tails that differ from the more normally distributed returns assumed for market VaR models. This is shown in Figure 23.3.

This difference in risk profiles does not prevent us from assessing risk on a comparable basis. Analytical method market VaR models consider a time horizon and estimate VaR across a distribution of estimated market outcomes. Credit VaR models similarly look to a horizon and construct a distribution of value given different estimated credit outcomes.

When credit risk is modelled, the two main measures of risk are:

- *Distribution of loss:* obtaining distributions of loss that may arise from the current portfolio. This considers the question of what the expected loss is for a given confidence level.
- *Identifying extreme or catastrophic outcomes*. This is addressed through the use of scenario analysis and concentration limits.

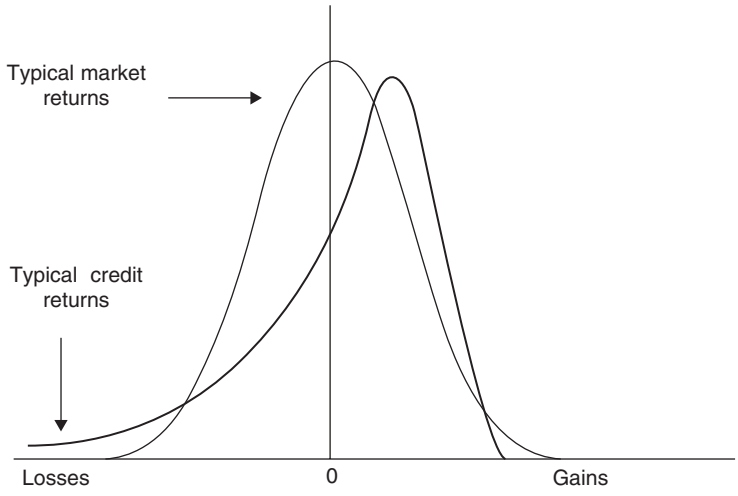


Figure 23.3 Comparison of distribution of market returns and credit returns

To simplify modelling no assumptions are made about the causes of default. Mathematical techniques used in the insurance industry are used to model the event of an obligor default.

Time horizon

The time horizon chosen will not be shorter than the time frame over which risk-mitigating actions can be taken. Credit Suisse First Boston (CSFB) (who introduced the CreditRisk+ model shortly after *CreditMetrics* was introduced) suggest two alternatives: a constant time horizon such as one year, or a hold-to-maturity time horizon. The constant time horizon is similar to the CreditMetrics approach, and also to that used for market risk measures. It is more suitable for trading desks. The hold-to-maturity approach is used by institutions such as portfolio managers.

Data inputs

Modelling credit risk requires certain data inputs. The common approach is:

- credit exposures
- obligor default rates
- obligor default rate volatilities
- recovery rates.

These data requirements present some difficulties. There is a lack of comprehensive default and correlation data, and assumptions need to be made at certain times. The most accessible data is compiled by the credit ratings agencies such as Moodys.

We now consider the CreditMetrics methodology for measuring credit value-at-risk.

CreditMetrics™

CreditMetrics was JP Morgan's portfolio model for analysing credit risk, providing an estimate of VaR due to credit events caused by upgrades, downgrades and default.

Methodology

There are two main frameworks in use for quantifying credit risk. One approach considers only two states: default and no default. This model constructs a binomial tree of default versus no default outcomes until maturity. This approach is shown in Figure 23.4.

The other approach, sometimes called the RAROC (risk-adjusted return on capital) approach, holds that risk is the observed volatility of corporate bond values within each credit rating category, maturity band and industry grouping. The idea is to track a benchmark corporate bond (or index) which has observable pricing. The resulting estimate of volatility of value is then used to proxy the volatility of the exposure (or portfolio) under analysis.

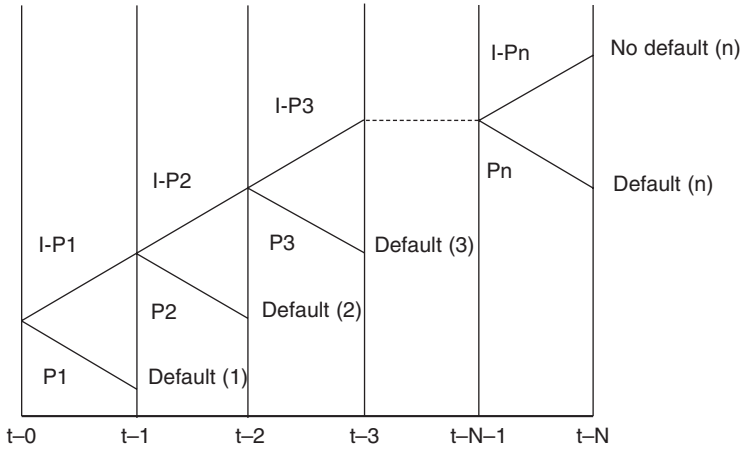


Figure 23.4 A binomial model of credit risk
Source: JP Morgan 1997.

The CreditMetrics methodology sits between these two approaches. The model estimates portfolio VaR at the risk horizon due to credit events that include upgrades and downgrades, rather than just defaults. Thus it adopts a mark-to-market framework. As shown in Figure 23.5, bonds within each credit rating category have volatility of value due to day-to-day credit spread fluctuations. The exhibit shows the loss distributions for bonds of varying credit quality. CreditMetrics assumes that all credit migrations have been realised, weighting each by a migration likelihood.

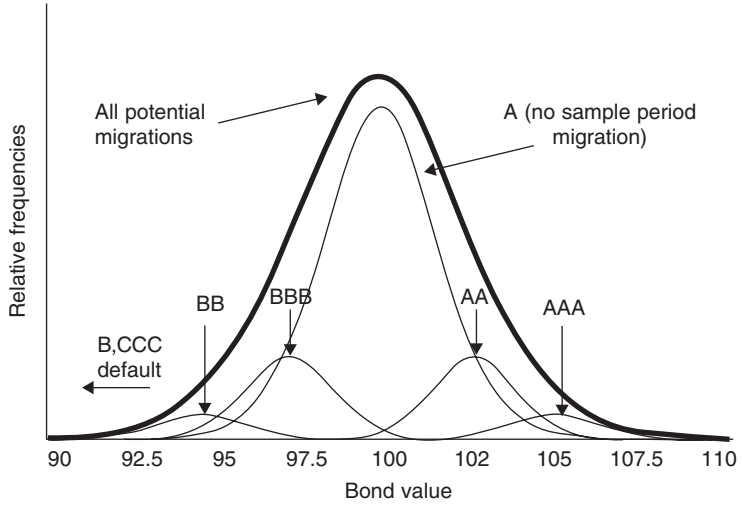


Figure 23.5 Distribution of credit returns by rating
Source: JP Morgan.

Time horizon

CreditMetrics adopts a one-year risk horizon. The justification given in its technical document (JPMorgan, 1997) is that this is because much academic and credit agency data is stated on an annual basis. This is a convenient convention similar to the use of annualised interest rates in the money markets. The risk horizon is adequate as long as it is not shorter than the time required to perform risk-mitigating actions. Users must therefore adopt their risk management and risk adjustments procedures with this in mind.

The steps involved in CreditMetrics measurement methodology are shown in Figure 23.6, described by JP Morgan as its analytical 'roadmap'.

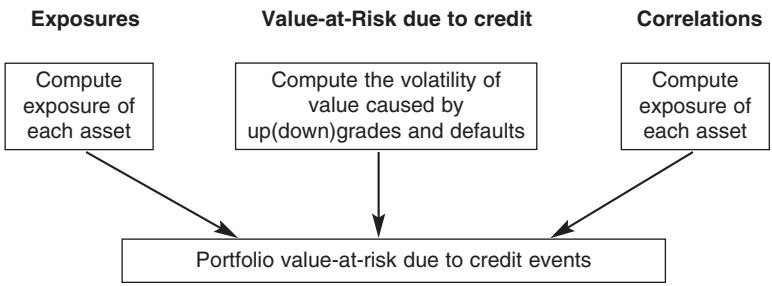


Figure 23.6 Analytics road map for CreditMetrics
Source: JPMorgan 1997.

The elements in each step are as follows.

EXPOSURES

User portfolio.
Market volatilities.
Exposure distributions.

VAR DUE TO CREDIT EVENTS

Credit rating.
Credit spreads.
Rating change likelihood.
Recovery rate in default.
Present value bond revaluation.
Standard deviation of value due to credit quality changes.

CORRELATIONS

Ratings series.
Models (such as correlations).
Joint credit rating changes.

Calculating the credit VaR

CreditMetrics methodology assesses individual and portfolio VaR due to credit in three steps:

Step 1: It establishes the exposure profile of each obligor in a portfolio.
Step 2: It computes the volatility in value of each instrument caused by possible upgrade, downgrade and default.
Step 3: Taking into account correlations between each of these events, it combines the volatility of the individual instruments to give an aggregate portfolio risk.

Step 1: Exposure profiles

CreditMetrics incorporates the exposure of instruments such as bonds (fixed or floating rate) as well as other loan commitments and market-driven instruments such as swaps. The exposure is stated on an equivalent basis for all products. Products covered include:

- receivables (or trade credit)
- bonds and loans
- loan commitments
- letters of credit
- market-driven instruments.

Step 2: Volatility of each exposure from up(down)grades and defaults

The levels of likelihood are attributed to each possible credit event of upgrade, downgrade and default. The probability that an obligor will change over a given time horizon to another rating is calculated. Each change (migration) results in an estimated change in value (derived from credit spread data, and in default, recovery rates). Each value outcome is weighted by its likelihood to create a distribution of value across each credit state, from which each asset's expected value and volatility (standard deviation) of value are calculated.

There are three steps to calculating the volatility of value in a credit exposure:

- The senior unsecured credit rating of the issuer determines the chance of either defaulting or migrating to any other possible credit quality state in the risk horizon.
- Revaluation at the risk horizon can be by either the seniority of the exposure, which determines its recovery rate in case of default, or the forward zero-coupon curve (spot curve) for each credit rating category, which determines the revaluation upon up(down)grade.
- The probabilities from the two steps above are combined to calculate volatility of value due to credit quality changes.

Step 3: Correlations

Individual value distributions for each exposure are combined to give a portfolio result. To calculate the portfolio value from the volatility of individual asset values

requires estimates of correlation in credit quality changes. CreditMetrics™ itself allows for different approaches to estimating correlations, including a simple constant correlation. This is because of frequent difficulty in obtaining directly observed credit quality correlations from historical data.

Example 23.3

An example of calculating the probability step is illustrated in Figure 23.7. The probabilities of all possible credit events on an instrument's value must be established first. Given this data the volatility of value due to credit quality changes for this one position can be calculated.

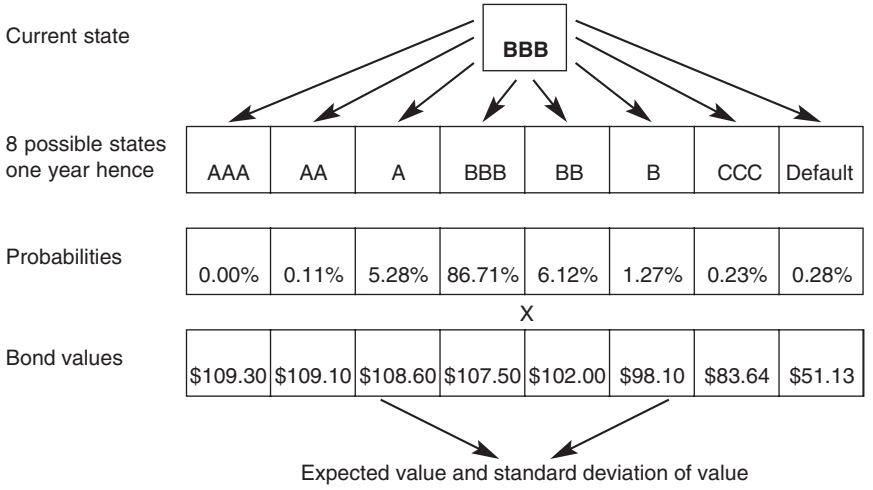


Figure 23.7 Constructing the distribution value for a BBB-rated bond
Source: JP Morgan 1997.

CreditManager™

CreditManager is the software implementation of CreditMetrics as developed by JP Morgan. It is a PC-based application that measures and analyses credit risk in a portfolio context. It measures the VaR exposure due to credit events across a portfolio, and also quantifies concentration risks and the benefits of diversification by incorporating correlations (following the methodology utilised by CreditMetrics). The CreditManager application provides a framework for portfolio credit risk management that can be implemented 'off-the-shelf' by virtually any institution. It uses the following:

- obligor credit quality database: details of obligor credit ratings, transition and default probabilities, industries and countries

- portfolio exposure database, containing exposure details for the following asset types: loans, bonds, letters of credit, total return swaps, credit default swaps, interest rate and currency swaps and other market instruments
- frequently updated market data: including yield curves, spreads, transition and default probabilities
- flexible risk analyses with user-defined parameters supporting VaR analysis, marginal risk, risk concentrations, event risk and correlation analysis
- stress testing scenarios, applying user-defined movements to correlations, spreads, recovery rates, transition and default probabilities
- customised reports and charts.

CreditManager data sources include Dow Jones, Moody's, Reuters, and Standard and Poor's. By using the software package, risk managers can analyse and manage credit portfolios based on virtually any variable, from the simplest end of the spectrum – single position or obligor – to more complex groupings containing a range of industry and country obligors and credit ratings. Generally this quantitative measure is employed as part of an overall risk management framework that retains traditional, qualitative methods.

CreditMetrics can be a useful tool for risk managers seeking to apply VaR methodology to credit risk. The model enables risk managers to apply portfolio theory and VaR methodology to credit risk. It has several applications, including prioritising and evaluating investment decisions, and perhaps most important, setting risk-based exposure limits. Ultimately the model's sponsors claim its use can aid maximising shareholder value based on risk-based capital allocation. This should then result in increased liquidity in credit markets, the use of a marking-to-market approach to credit positions, and closer interweaving of regulatory and economic capital.

Applications of credit VaR

Prioritising risk-reducing actions

One purpose of a risk management system is to direct and prioritise actions. When considering risk-mitigating actions there are various features of risk worth targeting, including obligors having:

- the largest absolute exposure
- the largest percentage level of risk (volatility)
- the largest absolute amount of risk.

A CreditMetrics-type methodology helps to identify these areas and allow the risk manager to prioritise risk-mitigating action.

Exposure limits

Within bank trading desks, credit risk limits are often based on intuitive, but arbitrary, exposure amounts. This is not a logical approach because resulting decisions

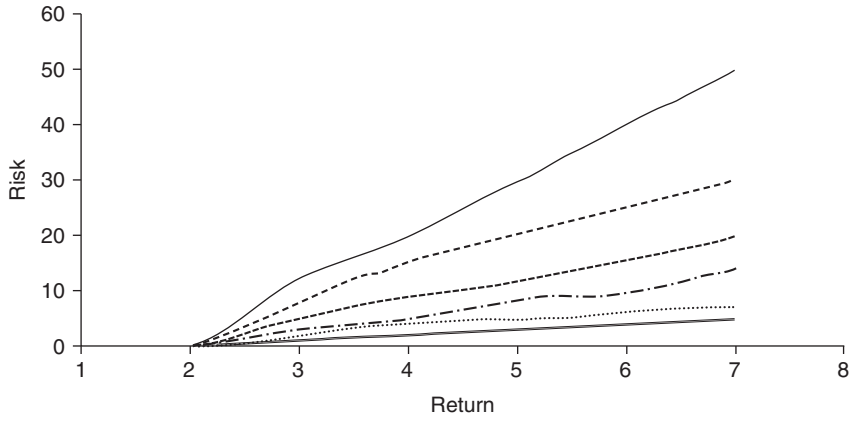


Figure 23.8 Size of total exposure to obligor-risk/return profile

are not risk-driven. Limits should ideally be set with the help of a quantitative analytical framework.

Risk statistics used as the basis of VaR methodology can be applied to limit setting. Ideally such a quantitative approach should be used as an aid to business judgement and not as a stand-alone limit-setting tool.

A credit committee considering limit setting can use several statistics such as marginal risk and standard deviation or percentile levels. Figure 23.8 illustrates how marginal risk statistics can be used to make credit limits sensitive to the trade-off between risk and return. The lines on Figure 23.8 represent risk/return trade-offs for different credit ratings, all the way from AAA to BBB. The diagram shows how marginal contribution to portfolio risk increases geometrically with exposure size of an individual obligor, noticeably so for weaker credits. To maintain a constant balance between risk and return proportionately more return is required with each increment of exposure to an individual obligor.

Standard credit limit setting

In order to equalise a firm's risk appetite between obligors as a means of diversifying its portfolio, a credit limit system could aim to have a large number of exposures with equal expected losses. The expected loss for each obligor can be calculated as:

$$\text{default rate} \times (\text{exposure amount} - \text{expected recovery})$$

This means that individual credit limits should be set at levels that are inversely proportional to the default rate corresponding to the obligor rating.

Concentration limits

Concentration limits identified by CreditRisk+ type methodologies have the effect of trying to limit the loss from identified scenarios and are used for managing 'tail' risk.

Integrating the credit risk and market risk functions

It is logical for banks to integrate credit risk and market risk management for the following reasons:

- the need for comparability between returns on market and credit risk
- the convergence of risk measurement methodologies
- the transactional interaction between credit and market risk
- the emergence of hybrid credit and market risk product structures.

The objective is for returns on capital to be comparable for businesses involved in credit and market risk, to aid strategic allocation of capital.

Example 23.4

Assume that at the time of annual planning a bank's lending manager says his department can make £5 million over the year if it can increase its loan book by £300 million, while the trading manager says it can also make £5 million if the position limits are increased by £20 million. Assuming that due to capital restriction only one option can be chosen, which should it be?

The ideal choice is the one giving the higher return on capital, but the bank needs to work out how much capital is required for each alternative. This is a quantitative issue that calls for the application of similar statistical and analytical methods to measure both credit and market risk, if one is comparing like with like.

With regard to the loan issue in the example above, the expected return is the mean of the distribution of possible returns. Since the revenue side of a loan (that is, the spread) is known with certainty, the area of concern is the expected credit loss rate. This is the mean of the distribution of possible loss rates, estimated from historic data based on losses experienced with similar quality credits.

In the context of market price risk, the common denominator measure of risk is volatility (the statistical standard deviation of the distribution of possible future price movements). To apply this to credit risk, the decision maker therefore needs to take into account the standard deviation of the distribution of possible future credit loss rates, thereby comparing like with like.

We have shown that as VaR was being adopted as a market risk measurement tool, the methodologies behind it were steadily applied to the next step along the risk continuum, that of credit risk. Recent market events, such as bank trading losses in emerging markets and the meltdown of the Long Term Capital Management hedge fund in summer 1998, have illustrated the interplay between credit risk and market risk. The ability to measure market and credit risk in an integrated model would allow for a more complete picture of the underlying risk exposure.

(We would add that adequate senior management understanding and awareness of a third type of risk – liquidity risk – would almost complete the risk measurement picture.)

Market risk VaR measures can adopt one of the different methodologies available; in all of them there is a requirement for the estimation of the distribution of portfolio returns at the end of a holding period. This distribution can be assumed to be normal, which allows for analytical solutions to be developed. The distribution may also be estimated using historical returns. Finally a Monte Carlo simulation can be used to create a distribution based on the assumption of certain stochastic processes for the underlying variables. The choice of methodology is often dependent on the characteristics of the underlying portfolio plus other factors. For example, risk managers may wish to consider the degree of leptokurtosis in the underlying asset returns distribution, the availability of historical data, or the need to specify a more sophisticated stochastic process for the underlying assets. The general consensus is that Monte Carlo simulation, while the most IT-intensive methodology, is the most flexible in terms of specifying an integrated market and credit model.

The preceding paragraphs in this section have shown that credit risk measurement models generally fall into two categories. The first category includes models that specify an underlying process for the default process. In these models, firms are assumed to move from one credit rating to another with specified probabilities. Default is one of the potential states that a firm could move to. The CreditMetrics® model is of this type. The second type of model requires the specification of a stochastic process for firm value. Here default occurs when the value of the firm reaches an externally specified barrier. In both models, when the firm reaches default the credit exposure is impacted by the recovery rate. Again, market consensus would seem to indicate that the second type of methodology, the firm value model, most easily allows for development of an integrated model that is linked through not only correlation but also the impact of common stochastic variables.

Appendix 23.1: Assumption of normality

The RiskMetrics™ assumption of conditional multivariate normality is open to criticism that financial series tend to produce 'fat tails' (leptokurtosis). That is, in reality there is a greater occurrence of non-normal returns than would be expected for a purely normal distribution. This is shown in Figure 23.9. There is evidence that fat tails are a problem for calculations. The RiskMetrics™ technical document defends its assumptions by pointing out that if volatility changes over time, there is a greater likelihood of incorrectly concluding that that the data is not normal when in fact it is. In fact conditional distribution models can generate data that possesses fat tails.

Higher moments of the normal distribution

The skewness of a price data series is measured in terms of the third moment about the mean of the distribution. If the distribution is symmetric, the skewness is zero.

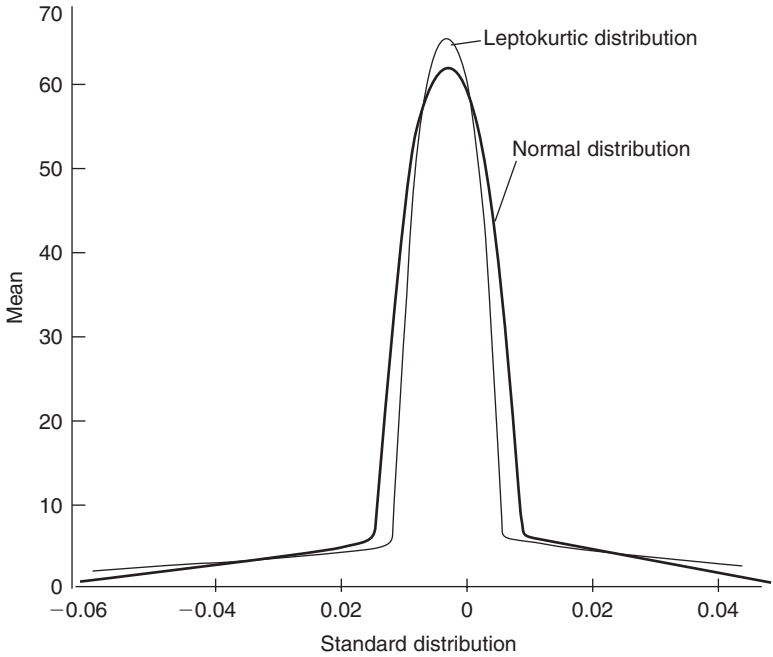


Figure 23.9 Leptokurtosis

The measure of skewness is given by

$$Skewness = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3/n}{\sigma^3} \tag{23.2}$$

The *kurtosis* describes the extent of the peak of a distribution, that is how peaked it is. It is measured by the fourth moment about the mean. A normal distribution has a kurtosis of three. The kurtosis is given by

$$Kurtosis = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4/n}{\sigma^4} \tag{23.3}$$

Distributions with a kurtosis higher than 3 are commonly observed in asset market prices and are called *leptokurtic*. A leptokurtic distribution has higher peaks and fatter tails than the normal distribution. A distribution with kurtosis lower than 3 is known as *platykurtic*.

Reference

JPMorgan, *Introduction to Credietrics™*, JPMorgan & Co., 1997.

PART VII
The 2007–9 Financial Market Crisis

The 2007–8 financial crisis, which precipitated a global recession, will be remembered as a milestone event in economic history. Although a proper study of it would belong in a textbook on economics or economic history, no standard textbook on finance can ignore it. Students and practitioners alike will need to familiarise themselves with the causes of the crisis, and its impact, which were widespread and carried on to 2009. The final chapter of this book considers the origins of the crisis and its impact across the financial markets. It concludes with policy recommendations for banks and regulators.

This page intentionally left blank

24
Origins and Impact of a Financial Crisis

> *Weakness has been contained to certain portions of the sub-prime market (and to a lesser extent, the Alt-A market), and is not likely to pose a serious systemic threat. Stress tests conducted by investment banks show that, even under scenarios of nationwide house price declines that are historically unprecedented, most investors with exposure to sub-prime mortgages through securitized structures will not face losses.*[1]

This was the assessment of the IMF in their quarterly report of April 2007. Only three months later one of the deepest crises since the Great Depression of the 1930s broke out. The IMF quote was characteristic for its time. Many people continued to believe that we were living in a new era where unlimited credit supply was possible without creating excesses. However the 'this-time-it-is-different' view appeared not to be so different. Mainstream economists either did not acknowledge the build up of the asset price bubble or underestimated its impact.

In this chapter we lay out the main causes of the 'Great Credit Crisis' of 2007–8. This is not as easy as it sounds. To blame only the US sub-prime mortgage market for this turmoil would be inaccurate and intellectually unsound. The Credit Crunch was the result of imbalances and excesses that had been built up over more than two decades. Before summing up these causes, we give an overview of the events of the crisis that started in mid-2007.

We then consider what factors triggered the crisis. We will explain how central banks have a certain responsibility in creating boom and bust cycles. The Clinton administration in the United States promoted social engineering via the housing market. The rise of sovereign wealth funds (SWF) and globalisation created a potentially destabilising ocean of liquidity. A 'shadow banking' system was set up by banks under the umbrella of 'special purpose vehicles', such as commercial paper conduits (ABCP) and Structured Investment Vehicles (SIVs). There were loopholes in the regulation of the banking and insurance industry. There was a build-up of systemic risk. The 'yen carry trade' was undertaken widely. There were

[1] IMF, *Global Financial Stability Report*, p. 7, April 2007.

serious mistakes made by bank senior management. The list of contributory factors is considerable: we begin with a chronological review.

Chronology of a crisis: overview of the great credit crisis 2007–8

If one were to draw a timeline, the starting point would be the summer of 2007. Major problems were coming to the surface in the US housing market. The quality of sub-prime loans was deteriorating rapidly and investors of home loan portfolios saw that they could not sell their loans in the secondary market any more, at any price. However, at the beginning of that year there had been signals that the mortgage market was having difficulties. In February 2007 HSBC and New Century, two of the major US lenders in the sub-prime market, hinted that they had to make sizeable provisions against their mortgage portfolio. Prices of sub-prime mortgage bonds, reflected in the ABX Index, had been falling systematically since July 2006. We see a sign of this at Figure 24.1, which shows the ABX-HE-BBB-06-01, an index used to create swaps based on 20 BBB-rated bonds sold in 2006 and consisting of home loans linked to borrowers with the lowest credit rating.

By April 2007 the mortgage market became seriously distressed, and New Century filed for Chapter 11 bankruptcy.

From then on events started to accelerate and on 4 May 2007 UBS AG had to close down one of its hedge funds, which was active in sub-prime, due to accumulated losses. Two weeks later one of the rating agencies put more than 60 tranches

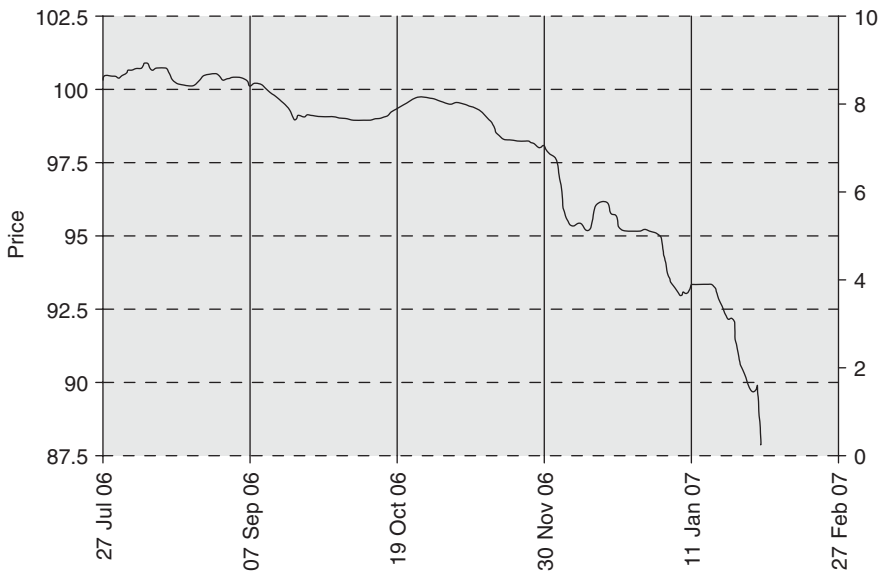


Figure 24.1 Rise of sub-prime mortgage bond risk: the ABX index
© Markit. Reproduced with permission

of 21 US sub-prime MBS deals on downgrade review which triggered further declines in prices on mortgage bonds.

There are many reasons why these loans started performing so badly, as we will explain later in this chapter. For one of the main reasons we need to go back to the beginning of 2001.

At the end of 2000 economic growth was contracting rapidly, causing the US Federal Reserve to cut interest rates aggressively. The events of 9/11 speeded up this process. By the end of 2001 rates were at 1.75% and fell to a 1% low (for that time) in 2003. This low interest rate environment made taking out a mortgage attractive to borrowers, as the housing market did not show any signs of slowing down despite the fact that the economy was faltering.

Many of these mortgages had a variable interest rate which would be revised annually or with a lock-in period of two or three years. However, from 2004 onwards, interest rates started rising back again to more normal levels. As Figure 24.2 indicates, by mid-2006 many homeowners saw the monthly payments on their mortgage being either doubled, tripled or even quadrupled – depending on when they had initially closed their mortgage – as either their lock-in period came to an end or their variable rate automatically increased with Fed fund rates. Simultaneously, health care costs, petrol and food prices had been rising rapidly, which reduced household disposable income.

Figure 24.3 shows that mortgage defaults were mainly concentrated around the loans that were closed with a variable interest rate at origination.

Figure 24.2 US Fed Fund rates: 2000–8
© Bloomberg L.P. Used with permission. Visit www.bloomberg.com

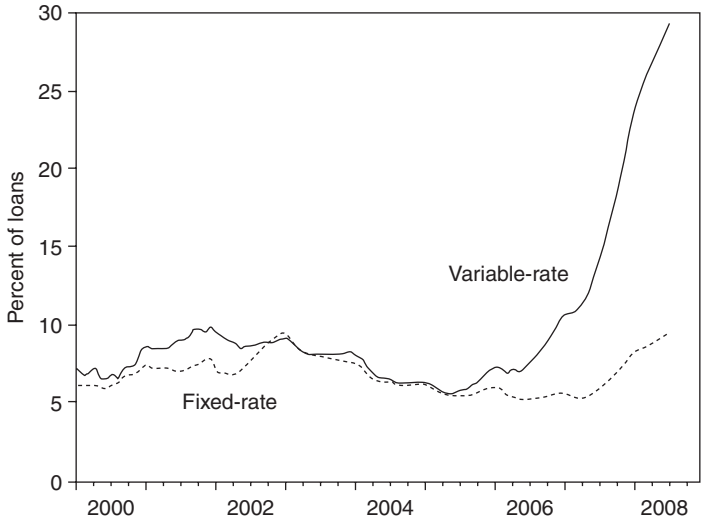


Figure 24.3 Sub-prime mortgage delinquency rates
Source: Federal Reserve 2008.

As interest rates rose, many homeowners found that they could not afford their monthly mortgage payments any more and started to default. These defaults were also the beginning of the end of the US housing boom, which created more problems for the mortgage market. Those who had borrowed 100% against the value of their house were confronted with negative equity (when the value of the house is lower then the outstanding debt). As a result banks that issued mortgages were facing reduced collateral guarantees.

This caused bankruptcies among US mortgage companies as well as the commercial banks that had purchased such loans from the mortgage lenders in the secondary market or via the securitisation market.

During the summer of 2007, banks became reluctant to continue lending to each other in the interbank market. Conditions in this market tightened considerably, and as a result the money market almost collapsed. The first casualty was the asset-backed commercial paper (ABCP) market, which virtually froze overnight. This had a knock-on effect on money market funds in serious trouble (see Figure 24.4).

This occurred because of the shadow banking system, which had been established over time, and was financing the purchase of structured credit securities via the issuance of short-dated ABCP. 'Asset-backed' CP was covered by a pool of loans, in some cases linked to the sub-prime market. The value of these loans was deteriorating rapidly and investors realised it was now risky to invest in ABCP, when previously the asset-backing feature of ABCP had served to make investors believe that ABCP was a safe investment. In addition to the ABCP market the Libor, repo and interbank market also stalled. In the case of repo the impact was understandable because it is also based upon secured lending.

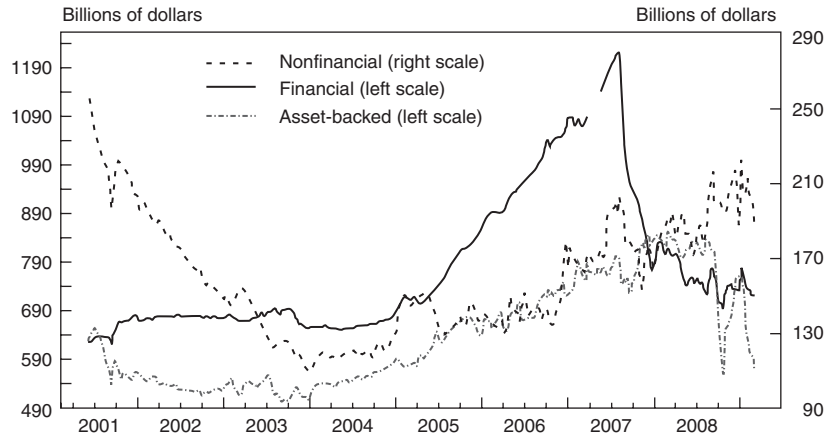


Figure 24.4 Asset-Backed Commercial Paper (ABCP) and non-ABCP market
Source: Federal Reserve Bank.

To illustrate how distressed the money market had become during the summer of 2007 we consider the TED-spread. The TED-spread measures the difference between the three-month USD Libor rate and the three-month US Treasury bill yield. The spread provides an indication of market perception of default risk in the interbank market. In Figure 24.5 we see the jump in spreads at the beginning of August 2007, during which two money market funds run by BNP Paribas reported serious losses. A similar spike is observed in September 2008 when Lehman Brothers filed for bankruptcy.

These series of events during summer 2007 were the beginning of the Great Credit Crisis and triggered what is sometimes called a 'Minsky moment'.

A Minsky moment can be described as the point in a credit or business cycle when investors start to have liquidity problems due to the spiralling debt they have built up in order to finance speculative investments. When this point is reached a sell-off wave starts, because no counterparty can be traced to make a bid at the high asking prices which were previously quoted. This results in a sudden collapse in market prices and a large drop in market liquidity.[2]

The US investment bank Bear Stearns became the first victim of this so-called Minsky moment. The bank was one of the highest leveraged investment banks on Wall Street and had lent substantial amounts to hedge funds over several years. After having bailed out two of their own hedge funds, one of their major clients – Carlyle Capital Corporation (CCC) a Dutch based hedge fund heavily invested in agency debt such as Fanny Mae and Freddy Mac paper – experienced difficulty following the receipt of margin calls from its lenders.

[2] Paul McCulley of PIMCO was the first to use this terminology when describing the Russian crisis in 1998.

Figure 24.5 TED Spread history, 2006–2009
© Bloomberg L.P. Used with permission. Visit www.bloomberg.com

As CCC could not meet these margin calls, their collateral was confiscated and this depressed the prices of agency debt further. However, Bear Stearns was known to be a large investor of this paper as well and the interbank market became suspicious about its ability to repay its money market liabilities. This affected the liquidity position of Bear Stearns substantially and the Federal Reserve had to come to its rescue with an emergency loan of $200 bln on 11 March 2008.

Unfortunately, the market considered this to be an indicator of serious problems and burgeoning systemic risk in the market. An e-mail from a hedge fund to Goldman Sachs asking to short Bear Stearns via the CDS-market was leaked into the market. Goldman Sachs refused to quote for the trade as it would increase their exposure towards Bear Stearns; this caused a run on the bank.[3]

After this point Bear Stearns could not attract liquidity in the interbank market and faced instant bankruptcy. An emergency rescue plan was worked out over the weekend of 14 March 2008 by the US government in cooperation with the Federal Reserve and JPMorgan Chase. The latter bought Bear Stearns at $2 a share or a total of $236.2 mio (this price was later on revised to $10 a share) with the backing of the Federal Reserve, which guaranteed Bear Stearns' trading obligations up to $30 bln. As we explain later in this chapter, this is an example of the Fed acting as a lender of last resort.

Although credit spreads tightened again and equity markets rallied on the back of this bailout, it soon became clear that the landscape of Wall Street was changing.

[3] See http://money.cnn.com/2008/03/28/magazines/fortune/boyd_bear.fortune/.

Despite the intervention of central banks across the globe to inject large amounts of cash into the banking system, liquidity became more and more of an issue. Investment banks (which cannot rely on a broad deposit base from their clients to fund their activities) found themselves restrained in the interbank market.

Liquidity dried up completely after the summer of 2008 as every major bank posted record losses for the last quarter and had to recapitalise urgently, since their Tier 1 capital ratios were not sufficiently high any more. Default risks were on the rise again and merchant banks such as Merrill Lynch, Morgan Stanley and even Goldman Sachs had to look for partners (commercial banks) as their wholesale funding model was no longer viable.

During September 2008 the major US investment banks were either merged with a commercial bank, such as Merrill with BoA, or reshaped into a commercial bank, such as Goldman Sachs, or went bankrupt, like Lehman Brothers.

In Figure 24.7 we observe the iTraxx Credit Index over the 2007–8 period, showing the points of big events such as the Bear Stearns bailout and Lehman bankruptcy against the credit spreads around the same time. Also notice the extreme low level of CDS spreads up until June 2007.[4]

As Figure 24.7 shows, from mid-2007 the credit market started to react with some sensitivity to these events. The equity market, on the other hand, displays a different picture. Compared to the credit market, equity prices did not drop substantially until Lehman Brothers filed for bankruptcy (see Figure 24.8).

In hindsight (although some commentators mentioned this at the time) it is clear that letting Lehman fail was the greatest mistake the US government, and Treasury Secretary Hank Paulson, made during the crisis. The 'hazardous put' was abruptly taken away from the market and caused serious fallouts for other counterparties. Lehman was an active player in the interest-rate and credit default swap market and its demise caused a serious increase in counterparty risk among other banks.

To illustrate what happened in the aftermath of the Lehman bankruptcy let us consider an example.

For example, Merril Lynch enters into a swap with one of their clients. Simultaneously Merril hedges itself in the interbank market with (say) Goldman Sachs. This is illustrated at Figure 24.9. In this situation all parties are fully hedged, and via multilateral netting agreements all exposures are eliminated. The problem is, however, that each party is only aware of its outstanding positions with its contractual counterparties; they may not know the full picture and therefore may not be aware of their indirect counterparty credit risk. If Merrill or Goldman do not want their clients to net their offsetting positions, both funds will have to either put up additional cash, or insure each other against counterparty credit risk by buying credit default swaps (see Figure 24.7, the iTraxx Chart).

This scenario unfolded shortly after Lehman's bankruptcy in September 2008. All the banks with counterparty exposure to it panicked that their counterparties would default, and they all bought credit default swap protection against each other. This explains partly why credit default swap spreads widened so excessively

[4] The chart shows the generic five-year CDS corporate spread on the iTraxx index.

Date	Description
…	…
02/04/2007	New Century – largest sub-prime broker –files for Chapter 11
23/06/2007	Bear Stearns bails out one of itsown hedge funds $3.2 bln
06/08/2007	American Home Mortage files for Chapter 11
08/08/2007	WestLB's ABS linked Asset Mgmt Fund suspends redemptions
09/08/2007	BNP freezes two Money Market Funds due to losses in loan market
14/08/2007	Goldman and clients inject $3 bln into their hedge fund
15/08/2007	Countrywide reports to beat risk for bankruptcy
17/08/2007	Fed cuts discount rate to 5.75% to ease credit crunch
14/09/2007	Northern Rock receives emergency loan from BoE
17/09/2007	UK Government guarantees deposits of Northern Rock
19/09/2007	Fed cuts Fed fund rates to 4.75%
24/10/2007	Merrill reports losson $ 8.4 bln write downs related to sub-prime
31/10/2007	Fed cuts Fed fund rates to 4.50%
27/11/2007	Citigroup to raise $7.5 bln capital from Abu Dhabi
10/12/2007	UBS reports $10 bln loss related to sub-prime
10/12/2007	Bank of America liquidates $12 bln cash fund
11/12/2007	Fed cuts Fed fund rates to 4.25%
13/12/2007	World central banks injecting $100 bln in inter bank markets
14/12/2007	Citigroup bailout its SIV with $ 58 blndebt
18/12/2007	ECB lends over $ 500 bln for Christmas
24/12/2007	Merrill raises $ 6.2 bln of capital
15/01/2008	Citiwrites down $ 18 bln
22/01/2008	Fed cuts Fed fund rates to 3.50% as markets implode
30/01/2008	Fed cuts Fed fund rates to 3.00%
14/02/2008	UBS reports $18.4 bln sub-primeloss
03/03/2008	HSBC reports $ 17 bln creditloss
07/03/2008	bac raises $ 1.5 bln to keep AAArating
14/03/2008	JP Morgan and Fed bail out Bear Stearns
01/04/2008	UBS writes down another $19 bln
08/04/2008	Washington Mutual raises $ 7 bln of capital
14/04/2008	Wachovia raises $ 5bln of capital
18/04/2008	Citigroup reports an additional $ 15 bln credit loss
22/04/2008	Merrill raises $ 9.55 bln of extra capital
22/04/2008	RBS announces to raise $ 24bln
29/04/2008	Citigroup raises an additional $ 3 bln of capital
29/04/2008	HBOS announces to raise $ 4bln of capital
06/05/2008	Fannie Maeraises $ 6 bln of capital
05/06/2008	MBIA and Ambac lose S&P AAA rating
09/06/2008	Lehman Brothers posts $ 3 bln loss and raises $ 6 bln of capital
10/07/2008	US Congress approves bailout plan for Fannie Maeand Freddie Mac
28/07/2008	Merrill writes down $ 5.7 bln and raises $ 8.5 bln new capital
07/09/2008	US Government takes over Fannie Mae and Freddie Mac
15/09/2008	Lehman Brothers files for bankruptcy
15/09/2008	BoA buys Merrill
16/09/2008	Fed lends $ 85 bln to AIG and take 80 %stake
18/09/2008	Lloyds TSB acquires HBOS for £12.2bln
25/09/2008	Washington Mutual seized by US government, assets sold to JP Morgan
29/09/2008	Iceland nationalises Glitnir Bank
03/10/2008	Dutch part of Fortis Bank gets nationalised
06/10/2008	Hypo Real Estate gets EUR 50 bln government bailout
13/10/2008	RBS, HBOS and Lloyds get $ 64 bln from UK government
13/10/2008	EU nations commit EUR 1.3 tln to bailout banks
16/10/2008	ECB gives Hungary EUR 5 bln credit line
22/10/2008	Wachovia reports $ 23.9bln loss
10/11/2008	Wells Fargo raises $ 11 bln to acquire Wachovia
19/11/2008	Citigroup buys $17.4 bln of SIV assets
15/12/2008	AIG sells $ 39.3 blnin assets to NY Fed's funds
20/12/2008	Latvia receives EUR 7.5bln loan from IMF and EU
23/12/2008	American Express to receive $ 3.39 bln in TARP funds
…	…

Figure 24.6 Chronology of major events in the Credit Crisis 2007–8
Source: www.creditwritedowns.com.

Figure 24.7 Behaviour of iTraxx index price during main events
© Bloomberg L.P. Used with permission. Visit www.bloomberg.com

Figure 24.8 Behaviour of S&P during main events
© Bloomberg L.P. Used with permission. Visit www.bloomberg.com

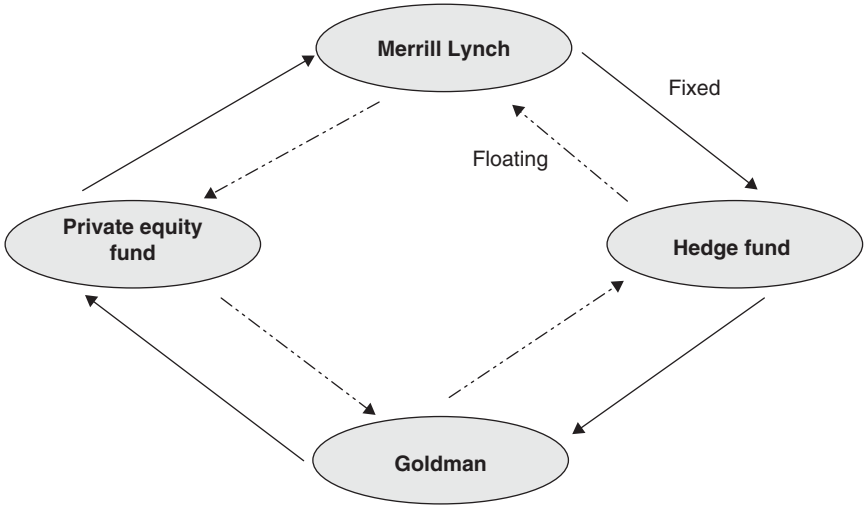


Figure 24.9 A network of interest-rate swap arrangements
Source: Markus K. Brunnermeier: Deciphering the Liquidity and Credit Crunch 2007–8. Used with permission.

in the aftermath of the Lehman Brothers collapse. A good example is the credit spread of AIG, which more than doubled within two trading days.[5]

From this point onwards, the Great Credit Crisis would spill over to the commercial industry – in other words from Wall Street to Main Street. As US Q4 2008 GDP numbers would later show, the global economy contracted sharply as commercial trade came to a standstill. Another effect of the tightening credit conditions would be banks' freezing of Letters of Credit (LoCs). LoCs are documents issued by a financial institution, presented to another financial institution to obtain release of goods (commodities, household articles, food, etc.). If banks refused to give you a LoC, you would simply have no economy.[6]

Figure 24.6 gives a chronological overview of the major events of the Great Credit Crisis from the summer of 2007 onwards. The list is certainly not exhaustive, as there were many more 'minor' incidents during that period.

We consider next what happened before the crisis, in other words the build-up of the Great Credit Crisis 2007–8.

The shadow banking system

One of the major causes of the Great Credit Crisis of 2007–8 was the build-up of a parallel off-balance sheet banking circuit next to the classic model of banking. This is known as the shadow banking system.[7]

[5] See Markus K. Brunnermeier, 'Deciphering the Liquidity and Credit Crunch 2007–2008', *Journal of Economic Perspectives*, Winter 2009.
[6] Gino Landuyt, 'Those Letter of Credits …', www.givanomics.com, January 2009.
[7] The term appears to first have been used by Paul McCulley and Bill Gross of PIMCO, in their weekly newsletters of November 2007. See www.pimco.com.

	1960	1970	1980	1990	1994	1999
Depository Institutions						
Banks	38.6	38.5	37.2	30.4	28.6	22.6
S&L (Mutual Savings)	19	19.4	19.6	12.5	7	4.3
Credit Unions	1.1	1.4	1.6	2	2	1.6
Total Depository Institutions	58.7	59.3	58.4	44.9	37.6	27.5
Insurance Companies	24	19.1	16	16.4	17.6	14.9
Pension Funds						
Private	6.4	8.4	12.5	14.9	16.2	18.7
Public (state and local)	3.3	4.6	4.9	6.7	8.4	11
Finance Companies	4.7	4.9	5.1	5.6	5.3	3.7
Mutual Funds						
Stock and Bond	2.9	3.6	1.7	5.9	10.8	17.2
Money Market	NA	NA	1.9	4.6	4.2	6
Total	100	100	100	100	100	100

Figure 24.10 Relative share of total financial intermediary assets (in %) US data
Source: US Board of Governors, Flow of Funds.

In a classic banking model a borrower would approach a bank for a loan. The bank would undertake credit analysis and decide to lend the client a certain amount at a certain spread over the rate it was paying to its deposit client basis. The deposit base is usually the major funding source of the bank's loan activities. At the end of the chain the deposits are backed by a lender of last resort (LoLR).

The LoLR is usually a central bank which has the ability to print money to support institutions that have liquidity difficulties. This usually at least delays or prevents a bank to go into bankruptcy or become insolvent, which in its turn would trigger sales and create M-t-M losses of any type of assets held on their books.[8]

During the late 1980s an observer would have seen a shift in the traditional role banks executed as their deposit base shrank and moved to mutual, pension and money market funds (see Figure 24.10).

It was not solely the loss of deposits that removed the banks' classic service model. From the 1970s onwards the CP and corporate bond markets developed in the US. Corporates became less dependent on banks for their funding, which they

[8] E. Philip Davis, 'Liquidity, Financial Crises and the Lender of Last Resort – How Much of a Departure is the Subprime Crisis?' in Paul Bloxham & Christopher Kent (eds), *Lessons from the Financial Turmoil of 2007 and 2008*, Reserve Bank of Australia, 2008, p. 115.

increasingly sourced directly from the market. As a result, banks started to look for alternative income streams and shifted towards a distribution-and-origination model – they started to package loans, took them off their balance sheets and sold them to capital market investors. This intermediary role produced fee-based income for banks.

Under this new model banks began to set up non-bank conduits such as structured investment vehicles (SIVs) in order to optimise their capital requirements under the Basel I and then Basel II regulatory frameworks. Via securitisation, the banks transferred the assets from their loan books to an off-balance sheet vehicle, such as a Collateralised Debt Obligation (CDO) or SIV. These vehicles had to look for alternative funding as they did not have the deposit base at their disposal that banks did. This issue was solved by financing the underlying liabilities via the sale of short-term paper such as ABCP, via medium-term notes (MTNs), or via the repo market. The average maturity of this paper was usually 90 days. As these vehicles did not have deposits backed by a strong lender of last resort, the only security line they had was a liquidity facility offered by the sponsoring bank in case they could not tap enough funds in the ABCP or repo market. This caused some serious liquidity risks for these banks, as on paper they would not be responsible for the loans which were securitised. However, indirectly – via this liquidity back-up line – they still were linked to these long dated liabilities, although this would not show up on their balance sheets.

Globalisation, sovereign wealth funds and the Greenspan conundrum

When searching for the reasons behind the current crisis we need to have a closer look at the role globalisation played. Globalisation was a powerful force in changing the landscape of financial markets. The seeds of globalisation were planted at the end of the 1970s. Before that, the US was more or less an autarkic economy that was focused on itself. Apart from international oil dependence the US economy was financed by its own pool of money.

The first major event was opening up financial markets in the US by broadening the investment guidelines of pension funds. They were allowed to invest in smaller mid-cap companies, which initiated the market in venture capitalism. 401k pension schemes freed up capital more, and by the mid-1980s during the Reagan Administration cross-border capital flows started to accelerate. The fall of the Berlin Wall in 1989 and the collapse of communism in general opened up trade opportunities across the globe and companies and banks started to operate more internationally.

One should also not underestimate the impact of the implosion of communism, as it freed up a substantial amount of capital which was allocated to defence spending during the Cold War. This 'peace dividend' contributed to a liberalisation of international trade as well, and increased productivity.

The banking industry observed opportunities in this new environment and expanded across borders and into new and emerging markets. This process was

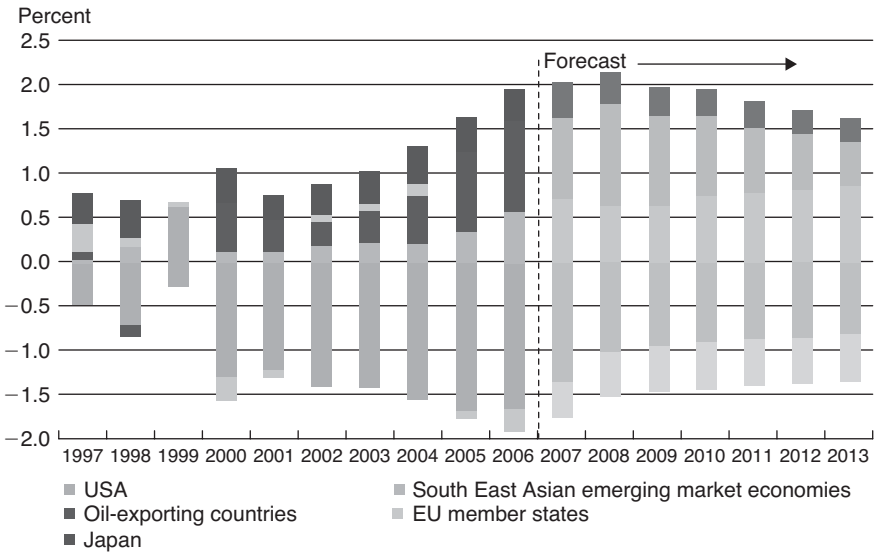


Figure 24.11 Development in US current-accounts balance as a percentage of GDP
Source: IMF World Economic Outlook, April 2008.

assisted by the development of technology and use of advanced IT infrastructure. For instance, transferring money electronically was now straightforward and improved bank liquidity.

President Bill Clinton further facilitated free trade by bringing down protectionist trade barriers. Globalisation flourished as markets opened up; new capital was made available to do business with Latin America, Asia, Central and East Europe.[9]

Over time these rising global flows of trade and capital also caused financial imbalances. The US started to build up a substantial current-account deficit, as the rest of the world was feeding rising US demand and consumption. Developing countries such as China and India participated in this commercial expansion and after a while southeast Asian emerging market economies and the oil-exporting countries were funding the US current-account deficit. This is shown at Figure 24.11.

A product of the Asian export-led growth economies and the oil-exporting countries was the Sovereign Wealth Fund (SWF), which managed the rising USD surpluses of these countries. SWFs are state-owned investment vehicles that invest their surpluses in global financial assets. Unlike central bank reserves, an SWF's portfolio is diversified in a wide range of assets such as equity, real estate, fixed income, hedge funds and private equity. Together with the hedge funds, Asian central banks and the private equity firms became very influential in the financial markets.[10]

[9] See David Smick, *The World is Curved*, Marshall Cavendish 2008.
[10] See McKinsey Global Institute, *The New Power Brokers: How Oil, Asia, Hedge Funds, and Private Equity are Shaping Global Capital Markets*, October 2007.

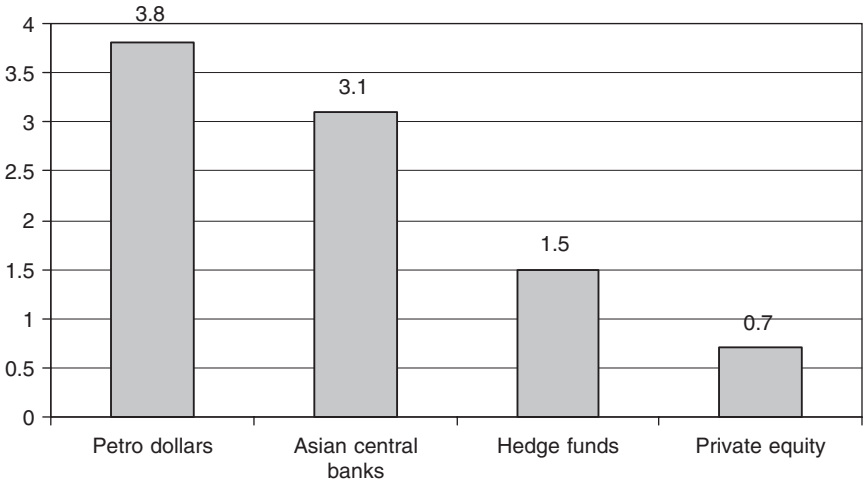


Figure 24.12 Assets under management in $ tln (2006)
Note: $1.5 tln of hedge funds are assets under management. Their real exposure is estimated to be leveraged up to $6 tln.
Source: McKinsey Global Institute 2006.

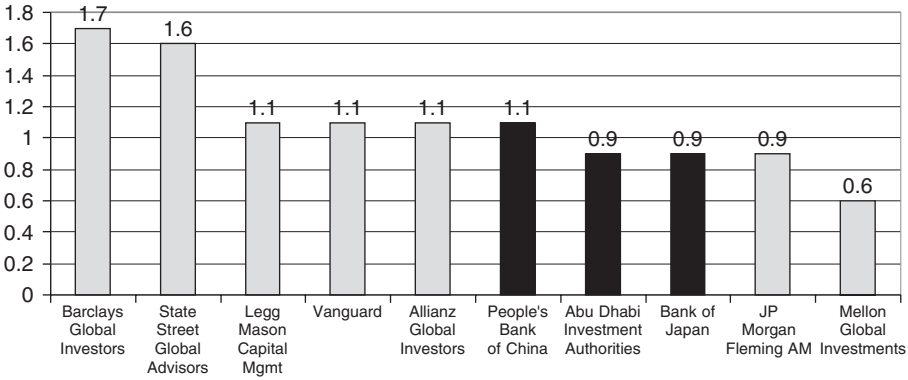


Figure 24.13 Top ten asset managers in $ tln (2006)
Source: McKinsey Global Institute 2006.

By 2006 the SWFs, together with the Asian central banks, became the biggest asset managers in the world (see Figures 24.12–24.13).

These new players added new liquidity to the global markets and by 2006 they represented (including the leverage part of hedge funds[11]) roughly $13.6 tln. Apart from the Asian central banks, the petrodollar countries were initially investing their reserves in US and European government bonds. This extra liquidity pushed

[11] The average leverage factor of the hedge fund industry as a whole is estimated at 4.

Figure 24.14 US 10-year Treasury Note yield 2000–9.
© Bloomberg L.P. Used with permission. Visit www.bloomberg.com

down long-term rates significantly. According to a McKinsey study, in the US bond market long-term interest rates were pushed down by an estimated 130 bp.[12]

This phenomenon is also known as the 'Interest Rate Conundrum' a term used for the first time by Mr. Alan Greenspan in June 2005. As showed in Figure 24.2, the Federal Reserve started hiking rates from 2004 onwards. However, despite aggressive hiking of short-term rates, the long-end of the USD curve continued to drop. This was a worldwide phenomenon. From June 2004 until June 2005 the US central bank hiked Fed funds approximately eight times from 1% up to 3%. Over that same period the yield on Treasury's benchmark 10-year note fell from around 4.8% to around 4% (see Figure 24.14).

At the time Mr. Alan Greenspan apparently was not aware, or did not believe, that this was due to large flows of foreign capital from export-growth countries running large USD surpluses and investing in USD assets. We know now that this was indeed the cause. As the price of oil almost doubled in value from 2002 to 2006, the oil-exporting countries in particular became one of the most important suppliers of capital (see Figure 24.15).

As new money continuously flooded into the market, this drove yields down, and investors became ever more demanding of yield. In addition there was a shift in risk appetite among SWFs and Asian central banks. Milton Friedman had

[12] McKinsey Global Institute, *ibid.*

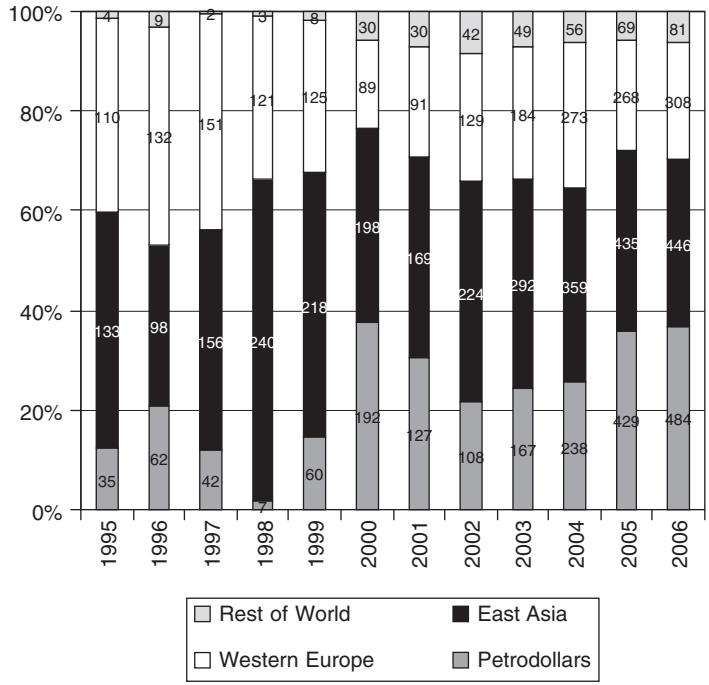


Figure 24.15 Net capital outflows from countries with current-account surpluses in $ bln
Source: IMF; McKinsey Global Institute.

warned many years previously that too much money chasing too few goods was dangerous, as it could result in asset-price inflation and bubbles – and this was indeed what resulted.

There is an argument to be made that the equity markets were driven more by corporate profitability and share buy-back programmes, while real-estate markets were clearly in a bubble. This was assisted by an unstated assumption among many investors and analysts that house prices would always rise. Private equity was boosted due to cheap credit available from the cash-rich banks. Cheap credit also boosted the hedge fund industry, which levered up using cash borrowed from banks. Finally, credit spreads were structurally pushed lower because of supply and demand, driven by synthetic CDO structures which were issued on a monthly basis by more and more western banks.

The Asian and oil-exporting countries played an important role in the dynamics of this prosperous environment.

The role of central banks

Central banks played a crucial role in the build-up of the Great Credit Crisis of 2007–8. In order to explain their responsibility, we need to have a closer look at their major tasks. One of these is keeping inflation under control. The Bank of

England (BoE), European Central Bank (ECB) and Bank of Japan (BoJ) all have explicit inflation targets, this being 2%. The US Federal Reserve (Fed) does not have an explicit inflation target, and is tasked with targeting growth.

> Trying to stabilise a consumer goods price index interferes with the allocation of resources over time.[13]

When the economy shows signs of overheating, central banks will decide to lower interest rates. By doing so, asset prices will start to rise simultaneously. Over the long run this will create asset bubbles.[14]

Due to inflation targeting, central banks have the tendency to overshoot their rate policy. This is exacerbated because much of their analysis focuses on lagging statistical indicators. Historically, it was felt that inflation could be controlled by controlling the money supply. The problem is, however – and there is an academic consensus around this issue – that a central bank has little impact on the money supply. From 2000 onwards the Fed stopped monitoring the M2 and M3 money supply statistics. Nevertheless, in times of economic slowdown, the Fed will still lower interest rates to stimulate the demand side. What is important however is that it takes away those stimuli in time, in order to prevent the economy from overheating. This is what happened after 2003.

Figure 24.2 illustrated the monetary policy action from the Fed from 2000 onwards. At the end of 2000 the US economy was losing momentum due to another bubble that was then deflating; the 'dot-com' bubble. During the 1990s we had seen an exponential rise in technological innovation. This arose partly from the changes the US economy had been going through from the late 1970s. Venture capitalism played an important role in stimulating research and development, which is the cornerstone of technological innovation.

The Internet and IT in general contributed to an acceleration in productivity growth. Among central bankers there was a growing consensus that due to this period of high innovation a 'New Economy' had been born: one which, due to faster productivity growth, could sustain this higher growth rate in an environment of continued low inflation. These were the dynamics of globalisation.[15] It was these forces that set in pace the build-up of two other crises earlier in the 1990s: the Asian currency and Russian debt-default crises (1997–8). This boom period was also driven by cheap money that was transferred to Asia and via currency manipulation, which led to a bubble in Asian economies.

When this cycle came abruptly to an end it put considerable downward pressure on commodity prices as local demand in Asian economies collapsed. Apart from the impact globalisation had on wages, this also contributed to a worldwide low inflation environment. In the meantime, asset prices kept on rising; but the Fed did not pay attention to this development. Following closely on the dot-com crash were the events of 9/11. With the example of Japan in mind, the Fed was concerned

[13] Quoted from Gerald O' Driscoll Jr., *Asset Bubbles and their Consequences*, Cato Institute, May 2008, p. 4.
[14] See also A. Hayek, *Prices and Production*, Routledge & Kegan Paul Ltd., 1935.
[15] See Alan S Blinder, *The Internet and the New Economy*, The Brookings Institution, June 2000.

about the risk of deflation in the aftermath of the terror attacks on the US and, even more importantly, the fallout from the burst of the dot-com bubble.

However, the downward pressure on consumer prices was a result of the forces of technological innovation and globalisation that were at play. Technological innovation had made it possible to produce more efficiently and bring down the cost of production. Globalisation opened up (labour) markets and had a downward pressure on wages. All this pushed global inflation down below the long-term average. In the meantime, also due to globalisation, the rapid rise of emerging markets such as China and India started to drive up the prices of food as many people saw their living standard improve significantly.

In hindsight, the Fed was clearly too late in hiking rates again, which may have stopped the housing bubble. As Mr. O' Driscoll, a former advisor of the Federal Reserve Bank of Dallas argues, a 'continued bias against inflation will produce a continued bias upward in price inflation'.[16]

The mismatch in timing of monetary rate policy versus economic growth is something central banks always struggle with. If one compares the change in Federal fund rates with the average growth rate in GDP terms during the previous two years one gets a clear view of the continuous overshooting of monetary policy by the Fed.

Figure 24.16 clearly shows that this connection was not a one-off incident. In 1974 the Fed cut rates aggressively from 10.5% to 5% over a period of two years. In the second year, though, the economy was already taking off again and the Fed wanted to catch up on their previous monetary easing; this contributed to the banking crisis of the 1980s. The central banks' intervention can be compared to a pendulum moving from one extreme to another.

The major problem here is the lack of focus towards asset price developments. As a matter of fact these are not reflected in the price indices the central bank follows. The BoJ failed to acknowledge the build-up of an asset bubble in Japan's economy during the 1980s. This triggered the 'lost decade' in Japan with deflation continuously hampering a sustained economic recovery. Mr. Greenspan gave a misleading sign to the market that, if an asset bubble were detected, the central bank would not do anything about it. The Fed would only deal with the fallout of the asset bubble.[17] This was later termed the Greenspan doctrine, or Greenspan put.

The Greenspan doctrine was nothing new; it only reiterated one of the reasons why central banks had initially been founded: to safeguard financial stability. The US central bank was formed in 1907 as an answer to a similar financial crisis to the Great Credit Crisis of 2007–8. In the build-up of that crisis, excessive credit expansion had been responsible for a run on the banking system. In order to prevent a repeat of a crisis of that proportion the US created the Federal Reserve system, which later became the Federal Reserve Bank we know today. Its main objective

[16]Gerald O' Driscoll Jr., *Asset Bubbles and their Consequences*, Cato Institute, May 2008.

[17]Speech by Alan Greenspan on 19 December 2002, Economic Club of NY: 'Asset bubbles cannot be detected and monetary policy ought not in any case to be used to offset them. The collapse of bubbles can be detected, however, and monetary policy ought to be used to offset the fallout'.

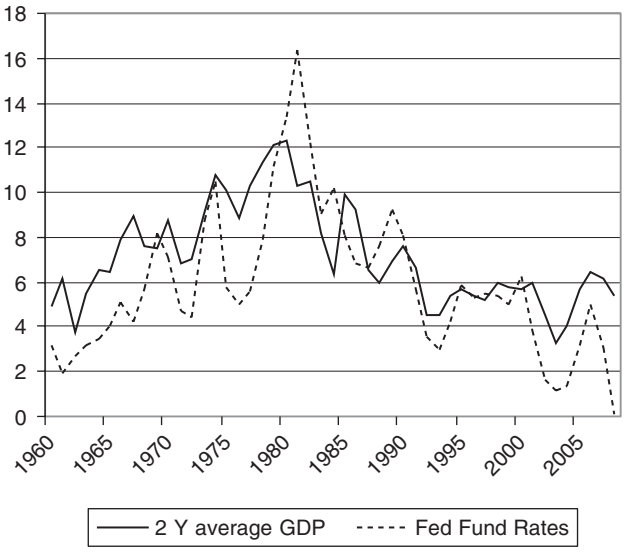


Figure 24.16 US two-year nominal GDP growth versus Fed fund rates 1960–2008
Note: Also see Brian S. Wesburry, 'A US Addiction to Easy Money', *Journal of Commerce*, October 1994 and Gerald O' Driscoll Jr., ibid.
Source: Bloomberg L.P.

was to act as a lender of last resort. In the event that any private bank faced a run on its deposits, the central bank would intervene to rescue the deposit holders.

The principle of the lender of last resort is dual. First of all, it provides strong comfort to deposit holders who are less likely to withdraw their money from a bank as the central bank has in theory put unlimited resources at the disposal of private banks to keep the credit process going. Secondly, it encourages deposit holders to place their money at the bank with the highest deposit interest rate.[18]

However, this high deposit rate results from a more aggressive lending policy. The principle of the lender of last resort is also known as *moral hazard*. What was initially set up to provide financial stability, indirectly contributes to boom and bust cycles. In times of crisis it brings stability to the system. However in the build-up of an economic cycle it drives the generation of lax credit practices. Banks that are more defensive in their lending policy are punished as they attract fewer customer deposits, and as a result they are competed out of the market.

As we indicated earlier, due to globalisation financial markets had become more interconnected. The Fed, under the supervision of Chairman Alan Greenspan, was aware of this development and expanded its role of lender of last resort for

[18]This summarises the Northern Rock case, in which deposit holders made the correct assessment; that the Bank of England would bail out the bank in case the credit portfolio should create substantial losses. Due to the more aggressive credit portfolio of Northern Rock, the bank was able to pay out a higher rate on clients' deposit accounts. See George Cooper, *The Origin of Financial Crises: Central banks, credit bubbles and the efficient market fallacy*, Harriman House 2008.

the US economy to an international level. During the Russian debt crisis the Fed preemptively cut rates in order to keep economic growth stable. This kind of intervention strengthens the perception that a central bank will come to the rescue if poor loan standards and irresponsible risk-taking by banks create a financial crisis.

In a speech at Wyoming in 2003 Mr. Greenspan said the following:

> 'At times, policy practitioners operating under a risk-management paradigm may be led to undertake actions intended to provide some insurance against the emergence of especially adverse outcomes. For example, following the Russian debt default in the fall of 1998, the Federal Open Market Committee eased policy despite our perception that the economy was expanding at a satisfactory pace and that, even without a policy initiative, was likely to continue to do so. We eased policy because we were concerned about the low-probability risk that the default might severely disrupt domestic and international financial markets, with outsized adverse feedback to the performance of the US economy'.[19,20]

In hindsight this was a mistake. It is apparent, then, that central banks take part of the blame for the Great Credit Crisis of 2007–8. Apart from the hazardous put and overshooting of monetary policy, another indication was already pointing to a major credit bubble due to fluctuations in inflation. The crisis was mainly credit driven (see Figure 24.17). Central banks consider themselves to bear the final

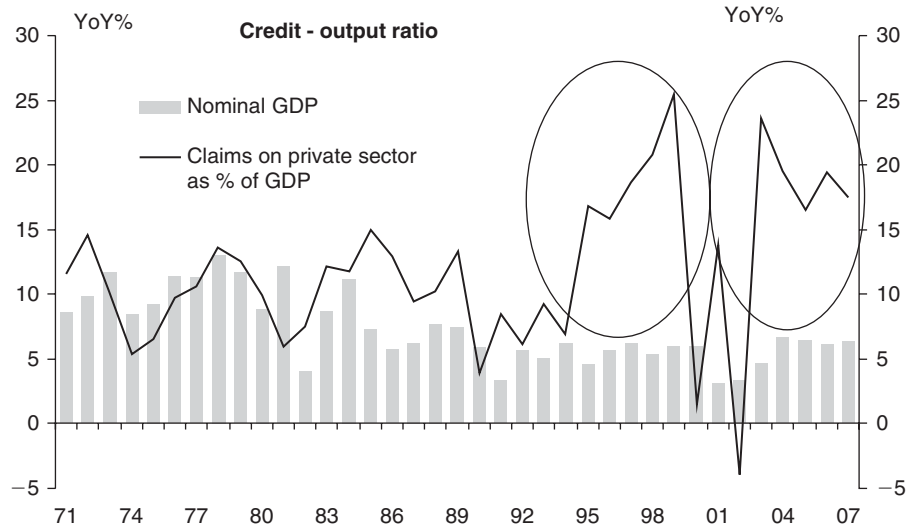


Figure 24.17 The US credit–output ratio
Source: ING March 2009.

[19]Alan Greenspan, *Monetary Policy under Uncertainty*, August 29 2003, Wyoming USA.
[20]George Cooper, *The Origin of Financial Crises: Central Banks, Credit Bubbles and the Efficient Market Fallacy*, Harriman House Ltd 2008.

responsibility of overlooking the credit and/or money supply. Unfortunately, they have little power in controlling this evolution. The money supply depends more upon the amount of leverage that the banking system is using, which is influenced by the degree of financial innovation.

By adjusting interest rates only, one has little to no impact on this development. To obtain control over these parameters one needs robust regulation.

The carry trade and rise of derivatives

During the 1990s Japan went through a deep recession. For more than a decade the BoJ had to fight against deflation by keeping short-term rates at virtually zero percent; and via quantitative easing – increasing the money supply and injecting new money into the banking system. This brought down the long-end of the interest rate curve.

Unlike the Great Credit Crisis of 2007–8, where deflationary pressures became a global problem, the risk of deflation in Japan remained an isolated issue. This permanent low interest rate environment created a situation where investors started to borrow in JPY and invest this elsewhere. This 'yen carry trade' appeared to be risk-free as long as the BoJ did not start raising interest rates and/or the JPY did not appreciate in value. However, the collapse of the hedge fund Long Term Capital Management delivered a serious warning about the potential consequences of a sudden substantial unwind of the carry trade.

Due to the large quantity of cheap credit available, and the tempting low interest rate conditions kept in place by the BoJ (but also the Fed), investors continued to set up a carry trade of unseen proportions.

Starting up a hedge fund became a lucrative business and substantial amounts were flowing into this 'new' industry.[21] By 2007 the total hedge fund community had approximately $3.1 trillions of assets under management including hedge fund of funds.[22] The hedge funds became one of the new power brokers in the financial markets due to their demand for banking services. On the other side of the trade were (investment) banks that were happy to lend to these new market participants, and in return received ancillary customer business for their capital market divisions.

The type of trades that were set up became ever more complex and the use of derivatives grew exponentially (see Figure 24.18).

Of course the carry trade was not available only in JPY. A carry trade can be set up in its own currency if the yield curve is steep enough. A steep yield curve is when short-term rates are much lower than long-term rates. As indicated earlier in the chapter the Fed created such a favourable environment for a considerable period of time. As long as the Fed did not hike aggressively and the yield curve remained steep, this would not become a problem for the investor, who would

[21]Although the first hedge fund dates from 1949.
[22]Data provided by Barclays Trading Group Ltd.

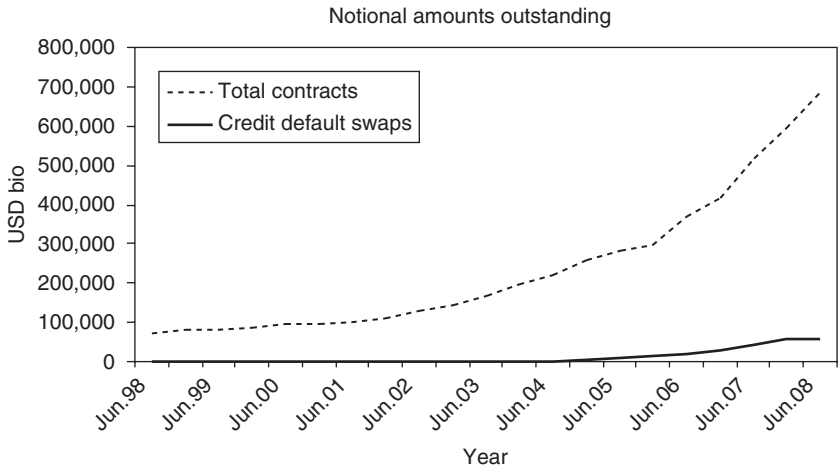


Figure 24.18 Growth of derivatives, 1998–2008
Source: BIS.

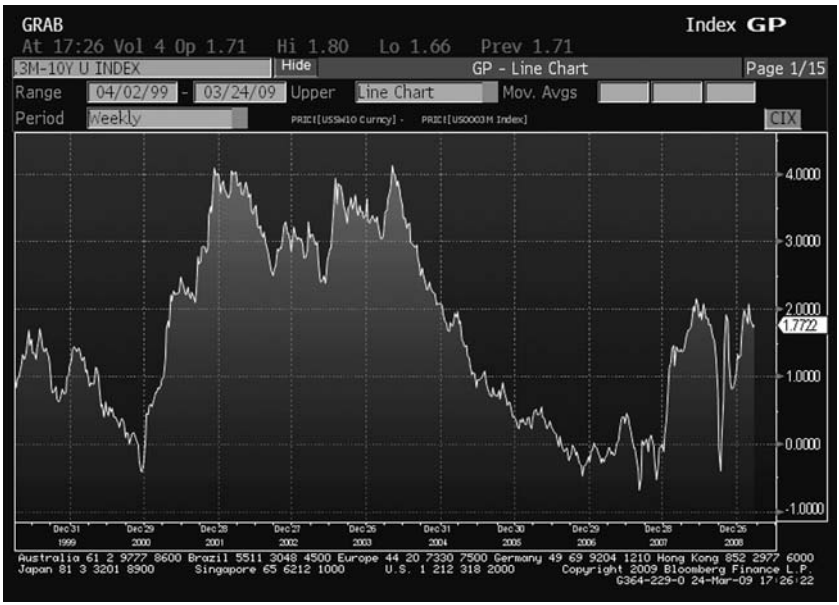


Figure 24.19 USD 3-month Libor versus USD 10-year swap rates 1999–2009
© Bloomberg L.P. Used with permission. Visit www.bloomberg.com

continue to borrow short-term money and invest this further down the curve in long-dated projects (see Figure 24.19).

This kind of trade was set up by US corporates, US banks, the US financial industry (hedge funds) and the US consumer. For instance,

- US government
  The US Government set up the carry trade by issuing the majority of its debt at below five-year maturity. One-third of its debt is even below one-year maturity. The purpose of this is obvious: keeping interest payments on the ever-rising US debt low. However, considering the large amount of outstanding debt, one runs the risk of an interest rate snowball effect. This mechanism is triggered when interest rates start rising to such levels that the government needs to start borrowing in order to pay off the interest payments of its debt.

  Some European countries such as Belgium and Italy faced a similar situation during the 1980s. They could only break through this vicious circle after devaluing their currency.
- US corporates
  In the same way as the US government, US corporates issued large amounts of short-term debt and invested it long-term via acquisitions. Corporate America always argued that this was not an issue as the balance sheets of US companies had improved since the last recession. This idea was only defensible as long as the stock market kept on rallying. A debt to equity ratio is only favourable when equity prices keep on rising.

  The US car industry, for instance, was issuing large amounts of short-dated debt and offering consumers finance rates close to zero percent to purchase cars.
- US financial industry
  Hedge funds and money centre banks set up the same trade. Depending on their risk profile they would invest in more risky assets such as emerging markets, junk bonds, mezzanine debt and the lower-rated and equity tranches of structured credit portfolios. Hedge funds would use a leverage factor by additional borrowing or via the use of derivatives.
- US banking
  The shadow banking system was only viable through the existence of low short-term interest rates. The SPVs of banks were funding huge amounts of CP with average maturities of 90 days or shorter. The money raised was invested in longer-dated assets of average maturity two years or longer.
- US consumer
  The US consumer was also setting up a carry trade via the housing market. People were buying houses via various inventive mortgage formulas. Adjustable rate mortgages (ARMs) are loans where the interest rate is linked to a short-term interest rate index such as the one-year constant maturity swap rate or a Libor rate. Two-thirds of sub-prime mortgage borrowers closed such ARMs.[23]

  Other mortgage products include: (i) bullet loans or interest-only mortgages, where the borrower does not make any down payment and could borrow 100% of the value of the property. The capital is only due at maturity of the loan. To make things even more aggressive this could be combined with the characteristics of ARMs, bringing the interest amounts usually to a minimum (as long as

[23]Noted in Jim Puplava, 'The Carry Trade Economy', *Financial Sense*, August 2004.

	Currency	Index Value Dec-08	Index Value Nov-08	Return Dec-08	Return Nov-08	Return YTD
Credit Suisse/Tremont Hedge Fund Index	USD	**351.08**	**351.2**	**−0.03%**	**−4.15%**	**−19.07%**
Convertible Arbitrage	USD	221.62	223.82	−0.98%	−1.88%	−31.59%
Dedicated Short Bias	USD	88.94	90.46	−1.68%	3.04%	14.87%
Emerging Markets	USD	264.49	263.92	0.22%	−1.87%	−30.41%
Equity Market Neutral	USD	225.47	224.54	0.41%	−40.85%	−40.32%
Event Driven	USD	395.52	400.56	−1.26%	−3.21%	−17.74%
Distressed	USD	452.18	463.96	−2.54%	−5.00%	−20.48%
Multi-Strategy	USD	371.03	372.86	−0.49%	−2.17%	−16.25%
Risk Arbitrage	USD	277.63	273.26	1.60%	−0.02%	−3.27%
Fixed Income Arbitrage	USD	166.79	168.13	−0.80%	−5.60%	−28.82%
Global Macro	USD	582.69	576.3	1.11%	1.54%	−4.62%
Long/Short Equity	USD	401.98	397.78	1.06%	−1.41%	−19.76%
Managed Futures	USD	284.19	277.61	2.37%	3.22%	18.33%

Figure 24.20 Credit Suisse/Tremont Hedge Fund Index Performance
Source: Credit Suisse/Tremont Hedge Fund Index.

short-term interest rates did not rise); and (ii) negative amortisation (neg-am) mortgages, in which some of the interest due is not paid when due, but is rolled up and added to the principle, so the borrower sees his/her outstanding amount actually increase over time.

The US consumer was the first to suffer when the carry trade was taken away step-by-step by the Fed as it raised interest rates from 2004 onwards. As short-term rates started to rise, the monthly payments on their mortgages started to rise significantly, until a point at which they started to default under the weight of their debt.

When the credit crisis unfolded hedge funds got into a vicious circle; as credit became tighter, it became more difficult to find funding to keep their model running. Credit spreads widened, and the mark-to-market of their underlying investments turned negative, which caused margin calls. This caused further unwinding of their positions, and turned into a vicious circle.

Even hedge funds that were running positive returns got into difficulties as investors were looking for liquidity and asked for early redemptions. The unwinding of certain carry trades, such as JPY borrowing, caused further extreme volatility, to the point that as markets crashed virtually no investment strategy worked: all of them suffered losses (see Figure 24.20).

The rise in derivatives use and the role of 'quantitative finance'

We highlighted earlier the rise in use of derivatives. This created extra liquidity and lowered barriers towards investors. Later we explain more about the role of securitisation, which was responsible for the spectacular growth in the use of credit derivatives. This growth did not stay limited to the credit market. Every market was involved: commodities, currency, equity and interest rates. The mortgage market was also an important driving force behind the use of interest-rate derivatives. It is not a surprise that the growth in the derivatives market went hand in hand with the rise in the housing market.

It was not only the housing market that contributed to this growth. As global trade under globalisation flourished, the need to hedge interest-rate and currency exposure rose exponentially too. Also, the hedge fund community was using complex trading and hedging techniques.

The complexity of derivatives increased significantly as there was blind trust in financial engineers and 'quants' who became the new wizards of the investment banking community. Universities started up new MBA and Financial engineering programmes in order to keep up with the demand from investment banks. These courses charged high fees and contributed to the cachet of the financial engineer as someone who could model market behaviour and accurately measure market risk using mathematical models.

This spectacular growth of the derivative business caused a serious problem: the build-up of systemic risk. By June 2008 the total notional amount outstanding was USD 683 trillion (see Figures 24.21 and 24.22).

Note that within the US almost 97% of all outstanding over-the-counter (OTC) derivatives are on the books of just five US banks. Almost 50% is on the books of only one single counterparty: JPMorgan Chase Bank.

This is not something new. In 1994 the US Government Accountability Office (GAO) reported on the risks of systemic failure. Its findings can be found on the GAO, and are noteworthy for their prescience.[24]

In essence five different risks arise when trading in derivatives. These are:

- Counterparty and/or credit risk:
  Entering into a derivative contract implies that you face a counterparty. From this moment you run the risk that your counterparty will not meet its obligations.
- Legal risk:
  As a derivative trade is a contract between two parties there is always the risk that a regulator or court will not accept the legal terms of the contract.

Type	Amount USD bln
Forex	62,983
Interest Rates	458,304
Equity	10,177
Commodities	13,229
Credit Default Swaps	57,325
Unallocated (*)	81,708
Total	683,726

Figure 24.21 The global OTC derivatives market
Note: (*) Includes foreign exchange, interest rate, equity, commodity and credit derivatives of non-reporting institutions, based on the latest Triennial Central Bank Survey of Foreign Exchange and Derivatives Market Activity
Source: BIS, June 2008.

[24] The full report is available at http://archive.gao.gov/t2pbat3/151647.pdf.

Rank	Bank Name	State	Total Assets	Total Derivatives	Total Futures (Exch Tr)	Total Options (Exch Tr)	Total Forwards (OTC)	Total Swaps (OTC)	Total Options (OTC)	Total Credit Derivatives (OTC)	Spot FX
1	JPMORGAN CHASE BANK NA	OH	$1,768,65	$87,688,008	$1,442,086	$2,349,629	$8,949,110	$54,385,247	$11,384,205	$9,177,731	$218,733
2	BANK OF AMERICA NA	NC	1,359,071	38,673,967	1,622,080	643,185	3,651,347	26,796,894	3,479,789	2,480,672	237,758
3	CITIBANK NATL ASSN	NV	1,207,007	35,645,429	253,586	432,226	5,071,607	20,210,646	6,737,581	2,939,783	536,543
4	WACHOVIA BANK NATL ASSN	NC	664,223	4,221,834	223,423	87,961	211,515	2,913,470	464,389	321,076	15,248
5	HSBC BANK USA NATL ASSN	DE	181,587	4,133,712	85,293	113,974	565,779	1,938,203	277,515	1,152,948	76,457
6	WELLS FARGO BANK NA	SD	514,853	1,429,088	174,358	21,694	468,891	562,659	199,766	1,720	19,149
7	BANK OF NEW YORK MELLON	NY	218,699	1,193,652	28,549	58,355	383,966	384,724	336,641	1,417	56,668
8	STATE STREET BANK & TRUST CO	MA	276,291	869,294	2,054	713	786,206	17,927	57,249	5,145	54,802
9	SUNTRUST BANK	GA	170,007	276,689	63,232	26,671	14,275	137,461	31,987	3,063	407
10	PNC BANK NATL ASSN	PA	134,780	198,478	26,441	12,500	6,079	124,859	23,660	4,940	1,580
11	NORTHERN TRUST CO	IL	68,930	175,128	0	0	165,238	9,232	389	269	22,761
12	KEYBANK NATL ASSN	OH	97,811	136,302	20,652	4,400	15,325	79,430	8,805	7,690	1,277
13	NATIONAL CITY BANK	OH	141,501	123,530	16,007	350	12,326	49,853	42,700	2,293	123
14	US BANK NATL ASSN	OH	242,597	97,056	1,640	9,000	23,871	51,272	9,618	1,655	878
15	MERRILL LYNCH BANK USA	UT	61,643	94,255	72,285	246	614	12,086	0	9,025	0
16	REGIONS BANK	AL	139,556	80,094	13,964	3,500	1,222	59,482	1,487	439	7
17	BRANCH BANKING AND TRUST CO	NC	133,166	71,044	3,599	0	8,632	49,228	9,533	52	57
18	RBS CITIZENS NATL ASSN	RI	132,609	59,474	0	0	4,890	53,129	1,228	228	37
19	FIFTH THIRD BANK	OH	67,318	58,101	94	0	8,999	39,367	9,333	308	863
20	LASALLE BANK NATL ASSN	IL	63,388	33,701	0	0	0	24,414	7,398	1,890	0
21	UNION BANK OF CALIFORNIA NA	CA	62,431	33,557	2,361	0	4,371	18,303	8,522	0	1,059
22	UBS BANK USA	UT	26,176	33,317	0	0	0	33,317	0	0	0
23	DEUTSCHE BANK TR CO	NY	43,932	27,004	0	0	391	20,941	601	5,071	0
24	MORGAN STANLEY BANK NA	UT	37,638	25,941	0	0	0	2,156	0	23,785	0
25	FIRST TENNESSEE BANK NA	TN	32,587	24,546	287	0	10,780	11,200	2,189	0	2
Top 25 Commercial Banks & TCs with Derivatives			$7,846,46	$175,403,202	$4,051,991	$3,764,404	$20,365,524	$107,985,498	$23,094,585	$16,141,200	$1,244,408
Other Commercial Banks & TCs with Derivatives			2,703,969	438,563	6,816	2,869	58,678	290,590	72,421	7,188	1,523
Total Commercial Banks & TCs with Derivatives			10,550,43	175,841,765	4,058,807	3,767,404	20,424,203	108,276,088	23,167,006	16,148,388	1,245,931

Figure 24.22 Concentration risk of OTC derivatives among US banks

- Systemic risk:
  The risk that financial markets will collapse due to the failure of one single market participant. This was almost the case with LTCM and was a real danger when Lehman Brothers collapsed.
- Market risk:
  The risk that the value of the underlying asset of the derivative moves in the opposite direction.
- Operational risk:
- Computer failure, human failure, and so on.

In addition to these risks, the GAO was also aware of potential developments that might create problems further down the road. As Figure 24.22 illustrates, there is a high concentration of derivative deals among only a few investment and commercial banks. This is reason enough to be worried about the potential danger of systemic risk.

Furthermore, due to globalisation financial markets have become interconnected over time. The first indication was the stock market crash of 1987, and the potential risk was even greater around the time of LTCM. The close ties between financial institutions are exacerbated by the high performance correlation between most asset classes. One lesson to be learnt from the Great Credit Crisis is that diversification in a financial meltdown will not protect one's investments.

Just a handful of banks dominate market-making activity in derivatives. This creates a high risk of disruption should one of them be forced to cease its treading business due to bankruptcy. As Martin Weiss notes:

> [T]he abrupt failure or withdrawal from trading of one of these dealers could undermine stability in several markets simultaneously, which could lead to a chain of market withdrawals, possible firm failures, and a systemic crisis.[25]

The earlier warnings were ignored, and after the collapse of Lehman Brothers the market received a serious warning about the impact of systemic risk. This issue remains unresolved.

Financial innovation and securitisation

In the previous section we noted the exponential rise of derivatives. This was only possible due to continuous financial innovation from investment banks. There was a strong belief among an (elite) group of academics that mathematical models could tame financial markets. All this was based upon the view that markets are or were rational, and would automatically correct excesses. However, we had seen earlier with the collapse of Long Term Capital Management that markets can get it wrong and that academic geniuses can make false judgement calls.

[25]Quoted from Martin D. Weiss, *Dangerous Unintended Consequences: How Banking Bailouts, Buyouts and Nationalization Can Only Prolong America's Second Great Depression and Weaken Any Subsequent Recovery*, National Press Club, March 2009.

It is worth pointing out here that the academic world is not purely driven by intellectual idealism but is a business industry in its own right. Universities started to anticipate the demand of Wall Street, training a new legion of rocket scientists with financial backgrounds – financial engineers and MBAs. This was a lucrative business; for example, a two-year MBA course might cost over $130,000 depending on the reputation of the university. As a result, investment banks recruited people they thought could outsmart the markets.

Risk can of course be passed on, but remains within the system. One can divide it among several people, but it will not disappear as some believed. This was the idea behind the hype of the CDO product, and the crucial mistake that re-insurance companies such as AIG made, when they guaranteed the super-senior tranches of structured credit securities.

This brings us to the cornerstone of the structured credit hysteria: securitisation.

In the whole process of globalisation – in which US consumers went on a spending spree buying goods from Asian and oil-exporting countries and funding this by taking equity out of their homes, and/or mortgage refinancing, and/or the build-up of the shadow banking system – securitisation played major role. Due to its characteristic of tranching a pool of loans into different risk categories, securitisation enables venture capitalists to participate in funding major projects, and this in the broadest sense of the word.

A very good example is the securitisation programme that Deutsche Bank launched in the microcredit market. Microcredit is very small loans that are extended to the unemployed, and extremely poor who lack collateral or don't have a credit history. This financial innovation originated in Bangladesh where it has successfully enabled extremely impoverished people to engage in self-employment projects that allow them to generate an income and, in many cases, begin to build wealth and exit poverty.

By securitising such a deal Deutsche Bank enabled private and ethical institutional investors to participate in the financing of this and make an effective contribution to the fight against poverty, distributing at least 120,000 very small loans to microbusinesses in 15 developing and emerging market countries.

There are many other examples that show how securitisation is a valuable tool and has considerable merit. Securitisation broadens the potential group of buyers and sellers due to its characteristics of diversification and customisation. As a result it increases liquidity and simultaneously reduces transaction costs. Mohamed El-Erian – Chief Investment Officer of PIMCO, the largest bond manager in the world – described these advantages very well:

> 'The combination of greater liquidity, diversification, customization is illustrative of the more general process of integration of financial markets. Like other components of this process, it is subject to excesses, including overproduction and overconsumption, but this need not necessarily derail the phenomenon over the medium term. Indeed, it may even render it more sustainable over the long term'.[26]

[26]Mohamed El-Erian, *When Markets Collide: Investment Strategies for the Age of Global Economic Change*, McGraw Hill 2008.

Securitisation has merits but also weaknesses. The problem is the disconnection between the arranger of the loans and the borrower itself. In classical banking the lender does its due diligence, runs credit tests and – depending on the risk – lends money with a certain credit spread on top of it. During the life of the loan the bank closely monitors the risk profile of the borrower. If the borrower experiences difficulty, the bank either renegotiates the deal, for example obtaining more collateral, or puts reserves aside for future potential write downs.

In many cases, with securitisation this risk-awareness disappears. The loans are packaged in more liquid securities and placed among investors. The bank that originally lends the money does not need to follow up the risk profile of the borrower any more, nor does the bank need to make reservations for future potential losses.

However, securitisation contributed greatly towards the high rise of globalisation. The technique made it possible to open up markets for investors who otherwise would never have the chance to invest in certain areas.

Political interference

There has always been tension between the financial industry and the political world. The major problem is that both worlds do not fully understand how they operate properly. Sometimes they find a consensus and good initiatives are taken; such as broadening the investment scope of pension funds, which triggered venture capitalism, or pushing through legislation which supported free markets and trade.

Sometimes mistakes are made, by imposing ill-conceived legislation upon the industry. The sub-prime crisis is an example of this. The seeds of the crisis were planted more than ten years before the turmoil started.

To look at this we need to go back to the early years of the US Clinton administration. One of President Clinton's main policy objectives during his first term in the White House was to promote home ownership for the weakest social classes. This policy is a noble one as it has many advantages. It is a way to build up savings among the poor, which will raise their living standards and will indirectly have a positive impact on crime figures. Other social advantages would be better education and a heightened sense of community.

The problem was that banks were very conservative in their lending policy towards this sector of the populace, as often they could not show a decent credit history or deposit a significant amount of cash to participate in the purchase of a house.

In order to work out his Housing programme President Clinton appointed Ms. Roberta Achtenberg as Secretary of Fair Housing and Equal Opportunity under supervision of the Department of Housing and Urban Development (HUD). Her main idea was that the refusal of mortgage lending towards the lowest social classes was based upon racial bias as the majority of this group were Hispanics or African Americans. As a consequence, she was convinced she could change the lending behaviour of banks by using anti-discrimination legislation.

A study conducted by the Federal Reserve Bank in Boston in 1992 gave support to this idea. The study showed that there was a biased mortgage lending towards Hispanics and African Americans which could be explained solely from a low credit quality perspective. It appeared that even at substantially higher interest rates, minorities were denied mortgages compared to whites.[27] The study was highly controversial as it later turned out that the data the Fed had used was of very poor quality. Nevertheless, the Boston Fed stuck to its conclusion.

Enforcement offices were set up all over the country, where people who had been turned down on their mortgage application could get counselling from public attorneys and investigators. Banks were put on trial, with million of dollars of claims based upon 'institutional racism and disparate treatment'.[28]

Banks started giving in under the pressure of legal claims, and offered mortgages with only 3% of deposits paid upfront. From 1994–9 home ownership among Hispanics and African Americans rose by two million.

Besides legal pressure, the Clinton administration changed the Community Reinvestment Act (CRA). Under this law banks would be subject to a rating system where the government would monitor how many mortgages they issued in low-income neighbourhoods. Banks needed to score well on this CRA rating in order to have a sign-off from the regulators on opening new branches and M&A activities.[29]

Next to Secretary Achtenberg there was Secretary Andrew Cuomo, overlooking the whole HUD. In this department Fannie Mae and Freddie Mac (also called government sponsored enterprises or GSEs) were directly under his supervision. Secretary Cuomo was using these two institutions as instruments to maximise home ownership among the lowest social classes of the US. Between 1997 and 2001 he took a series of initiatives which would push the GSEs into the sub-prime market. Fannie Mae and Freddie Mac do not actually sell or lend out mortgages. They buy these mortgages from banks and mortgage brokers. The purpose of buying up these loans is to free up the balance sheets of banks so that they get capital relief and can then originate more mortgages. These GSEs also have the mandate to buy mortgage-backed securities. Via this purchase mechanism the government can directly intervene in the housing market and inject money into the mortgage market.

The secretary of the HUD monitors this process. The aim of GSEs is to achieve the difficult balance between creating shareholder value – which would tend to make them more prudent in their lending policy – and meeting the goals to

[27]Stan J. Liebowitz 'Anatomy of a Train Wreck: Causes of the Mortgage Meltdown', *Independent Policy Reports*, The Independent Institute, Oct 2008.

[28]'Institutional racism: if a bank refused loans to proportionally more black applicants than white ones, the banks themselves would have to prove whether they had good reasons to refuse these loans'.

'Disparate impact: a form of discrimination as abstract and rarefied as to be imperceptible to its supposed victims, and often only discernible at all through the application of multivariate regression analysis to information stored on regulators' databases'. These are quoted from www.spectator.co.uk.

[29]See the feature on www.spectator.co.uk.

finance loans for those who need them. Every four years the Secretary of the HUD sets out new goals for the GSEs in terms of how many mortgages (and, more importantly, which type of mortgages) they need to buy up. Secretary Cuomo's predecessor, Mr. Henry Cisneros, started with imposing quotas on the GSEs but these were less aggressive and challenging than those Mr. Cuomo would later apply. Secretary Cuomo decided the GSEs would have to buy 50% of mortgages in very low-income neighbourhoods. He did not use the word sub-prime. In effect he was arguing that Fannie Mae and Freddie Mac were guilty of selective racial mortgage buying.

The targets which were imposed by Secretary Cuomo took effect from late 1996 onwards. The effect of the new policies was apparent as home ownership rose rapidly (see Figure 24.23).

The price the economy would pay later on turned out to be significant as well. Fannie Mae and Freddie Mac started to build out their sub-prime mortgage portfolio rapidly; as an indication, in 2000 Fannie Mae's sub-prime mortgages went from \$1.2 bln to \$9.2 bln and to \$15 bln the year after.[30] They also developed a new product that was called Alt-A. These were mortgages for borrowers with a slightly better credit history than the typical sub-prime borrowers, but who did not have to provide full information of their credit history.

In 2004 Freddie Mac warned the HUD that 'the increased goals created tension in its business practices between meeting the goals and conducting responsible lending practices'.[31]

The HUD has the responsibility to monitor the GSEs' targets. This is where things went wrong. Various consumer organisations had urged the HUD to monitor this 'housing for the poor programme' very closely as many were aware of the risks

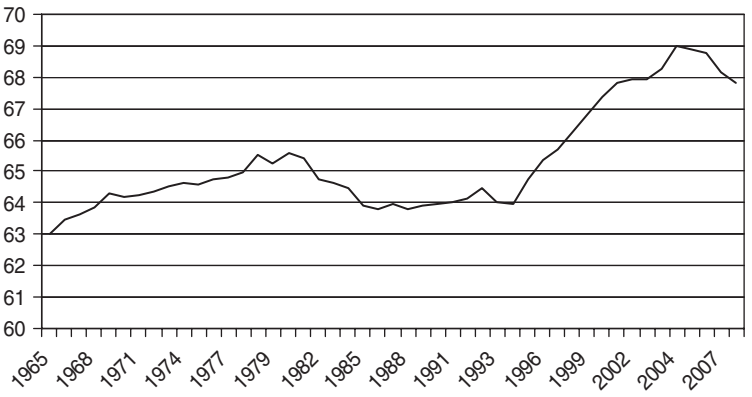


Figure 24.23 Percentages of US home ownership, 1965–2007
Source: Census.

[30]Source: http://www.villagevoice.com/content/printVersion/541234?ref=patrick.net.
[31]*Ibid*.

involved. The HUD decided it was not necessary to impose new or additional reporting rules under the revised GSE programme of 2000.

There is also the issue of the yield spread premium or kick-back payments the mortgage brokers received when they negotiated a new mortgage loan. Under this system a mortgage broker would not look for the cheapest mortgage available in the market, but would sell the borrower a more expensive one. When the mortgage broker sold the mortgage to one of the GSEs it would receive a side payment – which would be the difference between the rate the borrower was paying and the lowest rate available in the market.

It was Secretary Cuomo himself who issued a ruling in 1999, which courts applied from then on, saying there was nothing illegal in this practice.[32] Simply cancelling this practice would have prevented aggressive brokerage selling.

In this respect politicians, too, bear a huge responsibility in the crisis. As social engineering is a noble cause it becomes dangerous when an administration starts behaving ruthlessly in trying to change the world for the better.[33,34]

There is also discussion about whether or not the cancellation of the Glass-Steagall Act of 1933 was a catalyst for the credit crisis. The Glass-Steagall Act was created to prevent commercial banks becoming involved in investment banking and/or brokerage business. The conflict of interest between the two banking units was a major contributor to the Great Depression at the time. In 1999 Congress decided to lift the Glass-Steagall Act; giving US banks the opportunity to become a financial supermarket, and offering classical relationship banking together with underwriting business. Some argue that this gave the green light to banks such as Citigroup, Pain Webber (later a part of UBS) and Bankers Trust (the investment banking arm of Deutsche Bank) to start setting up SIVs and becoming active in the securitisation business.

It is highly unlikely though that this would have prevented the developments of securitisation, the CDO business and even the shadow banking system. Certainly, securitisation and CDOs already existed in the late 1980s, and were becoming more popular under the pressure of Congress to push for more mortgage loans to people with inferior credit profiles. Even without the cancelling of the Glass-Steagall Act SIVs would have been set up by investment banks to get maximum leverage. The major problem was that there was no regulation on leverage for banks in general.

Furthermore, the Glass-Steagall Act did not have a say in the European banking landscape and certainly would not have prevented (for instance) a small building society in Scotland from buying considerable amounts of this paper for so-called ALM purposes.[35]

---

[32]*Ibid*.

[33]www.marketorg.com.

[34]http://www.spectator.co.uk/the-magazine/features/2189196/part_6/clinton-democrats-are-to-blame-for-the-credit-crunch.thtml.

[35]Dunfermline Building Society, an institution whose business activities are concentrated solely in Scotland, defaulted and had to be partly nationalised by the Bank of England in 2009. It was found to be holding millions of dollars worth of US sub-prime ABS and CDO of ABS securities.

A final political interference was the prospect of continuous further tax cuts. One of the positive achievements of the Clinton administration was the rebalancing of the US household deficit. At the end of President Clinton's term the US Treasury was looking to create a structural (infinitive) budget surplus which could obtain huge accumulated amounts over time. Chairman Alan Greenspan also endorsed these tax cuts in front of Congress in order to temper the effect they might have had on the level of long-term interest rates. What was overlooked was that these surpluses were the result of extra tax income the government received by the bullish stock market and execution of stock option plans.

As soon as the dot-com bubble burst, any prospect of having structural surpluses on the US budget vanished as well.

Further regulatory shortcomings

Lack of regulation has always played an important role in boom and bust cycles. Usually there is a tendency towards deregulation at the end of a cycle.

The Savings and Loan Crisis during the 1980s was a good example of this. Although the S&L crisis was less complex than the Great Credit Crisis of 2007–8, it still had similar characteristics. If one replaced the shadow banking system and the housing market with high yield or junk bonds, one would get close to describing the S&L crisis.

Deregulation was the major factor at that time that triggered substantial losses for the US banking industry. Regulation Q put a limit on the interest rate that savings banks could offer their deposit client base. This gave them a competitive disadvantage in relation to the commercial banking industry, which brought them to the brink of closing down their offices. Basically, if there was no deposit base, one could not lend as a savings bank.

For that reason, Regulation Q was cancelled and savings banks put on an equal footing with commercial banks. The problem, however, was that commercial banks could lend to more risky projects and as a consequence could still offer higher deposit rates. Savings banks could not charge higher interest rates for residential mortgages, which was the basis of their lending income. In order to remain competitive with commercial banks in attracting cheap deposits, they started to invest in high-yield bonds offered by investment banks. This gave them a higher return which could justify a higher interest rate that they could offer their deposit client base.

Another point of similarity with the Great Credit Crisis of 2007–8 is that during the S&L Crisis, savings banks did not have a large number of credit analysts assessing the quality of these high-yield bonds.[36]

We have already explained how government interference and a lack of regulatory monitoring of the GSEs led to irresponsible buying of sub-prime mortgages and aggressive pitching from mortgage brokers. There were also other regulatory shortcomings which gave rise to the Great Credit Crisis of 2007–8.

[36]See Robert Barbera, *The Cost of Capitalism: Understanding Market Mayhem and Stabilizing our Economic Future*, McGraw-Hill Companies, 2009.

One example is the Basel I–II saga. Some blame Basel II for triggering the sub-prime crisis. There is a point to be made that Basel II reduced the risk weighting of sub-prime loans on the balance sheets of banks from 50% to 35%. Obviously this would give an incentive to banks to take more RMBSs onto their balance sheets. From 2004 onwards one clearly notices an acceleration of the RMBS market. There is a major problem, however, with saying that Basel II was responsible for this exponential rise. The decision to move from the Basel I framework to Basel II was taken in June 2004. The implementation of the Basel II framework only started from 2007 onwards. In the US it was implemented a year later, and was only mandatory for the largest 10 banks.

There was even a transition period ('parallel running') built in, to prevent major capital shocks on bank balance sheets. In the US those who had to comply with Basel II would leave everything unchanged until 2008. In 2009 they would move to 95%, followed by 90% in 2010 and 85% in 2011. Only from 2012 would a full application of Basel II be in place.[37]

It would be more logical to argue that the Basel I framework contributed to the build-up of the crisis. Basel I created the opportunity to build out a business model based upon SPVs, SIVs and CDOs.

Under Basel I, the rule for a bank was to keep 8% of capital against the risky loans they had on their balance sheet. But from the moment banks brought these loans into an SPV, they would put a credit line in place between them and the vehicle, and the capital they were required to place against the line was lower.

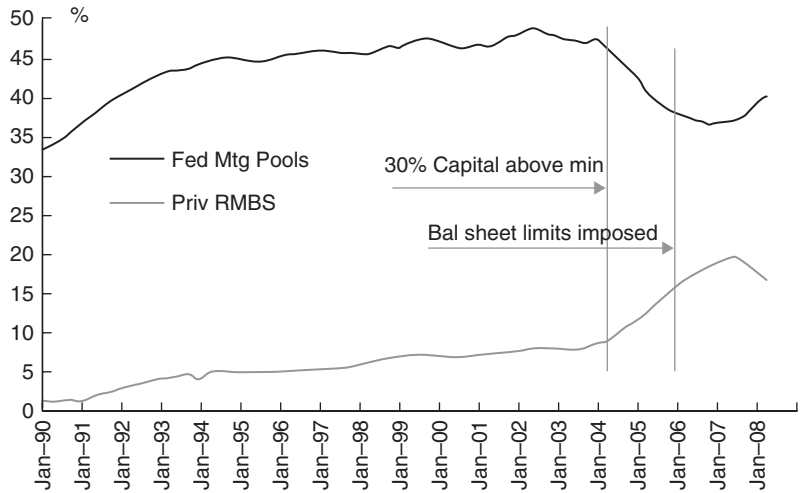


Figure 24.24 Fed mortgage pools versus private RMBS
Source: Federal Reserve, Datastream, OECD.

[37]See joint press release Federal Reserve, FDIC, Office of the Comptroller of the Currency of Thrift Supervision, PR-98-2005, September 2005.

This clearly underestimated the risk, as banks would still be linked to the outstanding loans via the credit lines that they had provided to these vehicles. This was the main incentive to set up a shadow banking system.

From the moment that the loans came back on their balance sheets (due to the liquidity problems their off-balance sheet vehicles were facing in 2007), the banks' capital ratio would be put under severe pressure. As explained earlier (see Figure 24.6: Chronology of major events in the Credit Crisis 2007–8), this is the very reason why a lot of banks had to come to the market during 2008 to raise additional capital.

Basel II made securitisation techniques less favourable for banks; however it remains to be seen whether it would have prevented the crisis. There were a number of problems which meant the dangers which caused the Great Credit Crisis 2007–8 were not addressed, such as the following:

- There is still a considerable dependence on internal quantitative models developed by banks to determine the regulatory capital. The quality of those models was heavily criticised during the crisis, as a lot of assumptions were made based upon very low historical correlation.
- It is still possible for banks to pool loans, place them in a structured vehicle and get capital relief. As Brunnermeier states (2009), 'Due to the reduction of idiosyncratic risk via diversification, assets issued by these vehicles will get a better rating than the individual loans in the pool'.[38]
- Basel II capital requirements encourage the pro-cyclical nature of banks' lending behaviour. Banks have to hold additional capital against greater anticipated losses as the economic cycle turns downwards. This could make an economic recession even deeper when banks are forced to restrict their provision of credit in a contracting environment.[39]
- Last but not least, Basel II is only applicable to European banks (and EU investment firms). It does not extend to non-bank financial institutions such as hedge funds and the large US brokerage houses.

As Basel I and II do not fully explain the acceleration in RMBSs, the root of the problem can be explained by a change in regulation on the GSEs in 2004. Adrian Blundell and Paul Atkinson describe this in a paper on the subprime crisis.[40] In the chapter on political interference we referred to the HUD monitoring the activities of Fannie Mae and Freddie Mac. This right was not exclusively reserved towards the HUD. The Board of Governors of the Federal Reserve, the FDIC and the Office of Federal Housing Enterprise Oversight (OFHEO) also have a regulatory mandate over the GSEs.

In 2004 the OFHEO raised the capital levels of Fannie Mae and Freddie Mac up to minimum 30% above the statutory minimum. This initiative was taken because of the significant build-up in sub-prime by both agencies in the years before, due

[38]Markus K Brunnermeier 'Decipering the Liquidity and Credit Crunch 2007–2008', *Journal of Economic Perspectives*, Volume 23, Number 1, Winter 2009.
[39]Rafael Repullo and Javier Suarez 'The Procycical Effects of Basel II', IMF, June 2008.
[40]Adrian Blundell-Wignall and Paul Atkinson, 'The Subprime Crisis: Causal Distortions and Regulatory Reform', paper presented at the Reserve Bank of Australia Conference, 2008.

to the pressure from Congress to increase home ownership. This sudden change of capital requirements had the same effect as the abrupt deleveraging process the market in general experienced from the summer of 2007 onwards when the credit crunch started.

A year afterwards, an additional restriction was imposed on the GSEs. The OFHEO introduced a balance sheet cap on both institutions; this after their balance sheet was already significantly limited due to the first ruling. If one of the agencies wanted to expand its activities above the balance sheet cap, they had to ask for approval on a case-by-case basis from the OFHEO. Until these two restrictions the GSEs were dominant players in the US mortgage market. Securitisation was already rising over the years but would accelerate after the intervention of the OFHEO. Before that, banks would simply sell their mortgages to Fannie Mae or Freddie Mac to get capital relief.

After the intervention, banks were facing considerable losses in revenues if this mortgage selling towards the GSEs was not replaced quickly. This pushed banks further into the securitisation business, as a new market suddenly appeared where no government competitors were active. Figure 24.24 clearly shows the contraction from the GSEs in 2004–5, and the simultaneous rise in private RMBSs issued.

In the United Kingdom, the local regulator (FSA) could take part of the blame as well, with its involvement in the collapse of Northern Rock. The FSA's decision to go ahead with the application of Basel II on Northern Rock's balance sheet triggered a textbook 'run-on-the-bank', which would result in a broad-based banking crisis in the UK.

Apart from the senior management of Northern Rock being far too aggressive in their lending policy – as one would expect from a small local player – the FSA as a regulator was aware of this activity. Giving the approval to start with the implementation of the new regulatory framework was a clear mistake, and the FSA must have been aware that this would cause some serious capital issues for the bank.

As Blundell–Wignall and Atkinson state,

> 'By June 2007, just as the crisis was to break and liquidity was to dry up, Northern Rock had total assets of £113 billion and shareholders' equity of £2.2 billion. Their RWA under Basel II was a mere £19 billion (16.7 per cent of total assets), compared to £34 billion under Basel I (30 per cent of assets). Under Basel II they had Tier 1 capital of a healthy 11.3 per cent of RWA, but only 2 per cent of total assets. When the crisis started, and liquidity dried up, they suffered the first run on a British bank since 1866, and their regulatory capital was less than 10 per cent of the £23 billion that the authorities used to support it'.[41]

The credit rating agencies

The expansion of the securitisation business and the creation of the shadow banking system was only possible with the assistance of the rating agencies. They

[41]See Blundell–Wignall and Atkinson, *ibid* page 78.

played a crucial role in the build-up to the crisis by issuing excessively optimistic ratings for structured credit paper.

They made a crucial mistake in the development of their default forecasting models. Mortgage default rates and delinquencies were all based upon historically low levels.

Even more worrisome was the assumption that real estate prices could only drop regionally but not across the US in general. Historical data never gave any indication of this phenomenon.

Following this assumption made correlation risk of real estate prices between regions in the US rather low. As a result this had a positive impact on giving a AAA value to the top tranches of a RMBS which would help the sale argument of the underlying bond. Blundell-Wignall and Atkinson make the argument that credit agencies were stimulated in this type of business as the fees linked to this type of bonds were much higher compared to the fees they received on rating corporate bonds. This concept is called 'rating at the edge'. We do not subscribe to this belief however.

On the other side there were the investors who were desperate for yield (the interest rate conundrum) and at the same time intrigued by the higher yields these structured credit notes offered, given the rating and low default probability.[42]

Bad management practices

We have already referred to the aggressive lending practices of the senior management of Northern Rock. One may wonder whether a small local player should get involved in complex and risky investment banking practises.

There was hype among banking executives to target the highest deals and market share expansion, without keeping the funding consequences in check. The authors of this book themselves experienced in their day-to-day activity the frustration of this Return on Equity (RoE) obsession. From around 2000 onwards, decent, ultra-conservative deals with an RoE of 10–12% were turned down on a regular basis as senior management would only start considering proposals offering an RoE of above 20%. These high-return trades later unravelled in the credit crisis at great cost.

This kind of development started to run a life on its own; and sooner or later this would result in a ponzi scheme as described by Minsky.[43] Market share and rankings were topics being discussed during weekly board room meetings. Expensive consultancy firms were hired to investigate how banks could close the gap with their competitors and how revenue could be driven up further. It is ironic that a term ('farmers-and-hunters') that was invented by the consultancy firm

[42]Our sources here are Joshua D. Coval, Jakub Jurek and Erik Stafford, 'The Economics of Structured Finance', Harvard Business School, *Working Paper No. 09-060*, October 2008, and Adrian Blundell-Wignall and Paul Atkinson, 'The Subprime Crisis: Causal Distortions and Regulatory Reform', paper presented at the Reserve Bank of Australia Conference, 2008.
[43]Hyman P. Minsky 'Stabilizing an Unstable Economy', McGraw-Hill Companies, 2008.

McKinsey became a dominant factor in the annual review of sales desk personnel in dealing rooms.

In order to realise these objectives, serious pressure was put on Treasury desks to offer the cheapest funding possible but without taking into account the liquidity consequences.

Remuneration of personnel did not consider the quality of profit and loss (P&L) created by either the trader or salesperson. This opened the door for excesses on fixed income desks. Some of them would exclusively focus on the sales of CDOs and hardly pay attention to the daily plain vanilla business of their client base, as the commissions on these deals were much more lucrative than quoting for a plain vanilla interest-rate swap. This took on outrageous proportions, that at some European banks the clients of the CP-desk were poached to buy CDO AAA-tranches as an alternative to their CP liquidity business. This verges on 'mis-selling' to the customer, who may not have been sophisticated enough to analyse what they were buying.

Apart from RoE, market share was a major concern and this led to a consolidation wave in the banking industry. For example:

- UBS's investment bank unit was encouraged to grow market share levels to approach that of Goldman and Merrill Lynch;
- Smaller European commercial banks moved up in the rankings by central and eastern European expansion;
- Tier 2 commercial banks tried to become Tier 1 banks by taking over competitors.

This expansion wave ultimately led to the creation of banks that, instead of 'too big to fail', became 'too big to rescue'. UBS's balance sheet grew to a point at which it far exceeded the GDP of Switzerland. A similar phenomenon took place at Fortis (which, together with Royal Bank of Scotland and Banco Santander, bought ABN-Amro but did not arrange a decent funding strategy to close the deal). As the credit crunch was at its height, the bank came to the brink of collapse and eventually had to be split up between the Dutch, Belgian and Luxemburg governments in order to make sure the deposit holders would not lose their money. Last but not least, there were the Icelandic banks that grew to astronomical levels driven by irresponsible lending standards, and drove their local government to virtual bankruptcy until the IMF itself had to come to the rescue.

It is hard to deny that all this was driven by incompetence, megalomania, arrogance, pride, machismo and greed; and by an utterly contemptible disregard for genuine shareholder value creation.

Conclusion and lessons to be learnt

It becomes more obvious that the Great Credit Crisis of 2007–8 was not exclusively caused by irresponsible sub-prime lending and a housing correction in the US.

More elements were at play, and the seeds of the crisis were sown over a prolonged period extending back to the 1980s.

The various causes that were laid out in this chapter do not have equal weight; it is the mixture or combination of these developments put together over a period of 25–30 years that created an enormous bubble that was due to burst at some point. Some elements, though, played a more detrimental role than others.

In Figure 24.25 we try to give a weighting to each of the single causes that contributed to the crisis, on a scale from one to five: one being very strong, two strong, three average, four small, and five very small.

The Lehman Brothers failure is also mentioned on the list. This was more an event that occurred during the unfolding of the crisis. However, since the consequences of systemic risk emerged in the aftermath of the bankruptcy of the bank, it became very clear that this was the biggest mistake the US government made in managing the crisis. Preventing the fall of Lehman Brothers and working out a similar solution – as Treasury Secretary Paulson did for Bear Stearns earlier that year, in cooperation with the Fed – would have kept the crisis more contained.

Putting together a list with different ratings is not an easy exercise, either, as these element are interconnected. For example, the shadow banking system could only be set up under the soft treatment of credit liquidity lines towards structured credit vehicles by Basel I and the use of securitisation.

Securitisation was a valuable tool which created added value and contributed to growth and opening up (financial) markets. However, the tool was abused; used for the wrong reasons. The credit rating agencies made clear mistakes in assessing the risk of pooled loans compared to loans on an individual basis. Therefore, we

Causes of the Great Credit Crisis 2007–2008	Approximate Start	Weighting
Sub-prime market/housing market	1992	2
Shadow banking system	2000	1
Globalisation and SWF's	1995	3
Central banks	1980	2
Greenspan's moral hazard	2001–2003	1
Systemic risk	2004	3
US consumer debt build up	1995	3
Carry trade	1987	3
Financial innovation and securitisation	1985	3
Rating agencies	2000	2
Lehman brothers bankruptcy	2008	2
Political interference	1992–1994	1
Regulation: GSE's	1994	1
Regulation: Basel I	1992	2
Regulation: Basel II	2004	4

Figure 24.25 Weighting of causes of the Great Credit Crisis 2007–8

give a higher rating to the contribution of the rating agencies than to the tool of securitisation itself.

A similar logic is used on the combination of central banks, Mr.Greenspan, the carry trade and the rapid growth of the US consumer debt. The latter is only the outcome of an opportunity created by the central banks. Central banks also bear some responsibility as they are the prime observers of money and credit supply, which clearly showed signs of running ahead of itself from the late 1990s onwards. The reiteration of the moral hazard by former Federal Reserve Chairman Greenspan in 2002, which was later called the Greenspan put, gave a free lunch to excessive risk taking. In this respect we consider Mr. Greenspan's verbal interventions even more severe than the workings of central banks.

Last but not least, the political pressure from the Clinton administration and the lack of regulation of Fannie Mae and Freddie Mac were major seeds which ultimately translated into a credit bubble.

In order to prevent a crisis of similar proportions in the future, it is important to acknowledge all these elements that contributed one after another to the Great Credit Crisis 2007–8. Going forward it will be important that certain adjustments are made to the financial system as we knew it for so many years.

On 1 April 2009, a G20 Summit in London was dedicated to the crisis and measures were taken to prevent a similar crisis from happening again. The major initiatives that were taken were:

- Tougher control and regulation of financial institutions
- Regulation of hedge funds
- Sanctions against tax havens
- Specific regulation to prevent systemic risk including the limitation of leverage
- Regulation to curb boom and bust cycles
- Stricter control of credit rating agencies.

These measures still needed to be worked out in detail during the printing of this book. Our recommendations on the commitments from the G20 meeting at that time were:

- Tougher control and regulation on financial institutions:
  Stricter regulation, so that commercial or retail driven banks are restricted from building a shadow banking system and jeopardising the deposit basis of their clients. Basel II still leaves the door open to restart this business.

  Much higher-risk weighting measures for commercial or retail driven banks to buy structured credit securities for ALM purposes. This would prevent a situation in which a small building society might buy CDO tranches to fine tune its ALM positions.

  It would also be useful to put a cap on the amount of a mortgage that can be offered as a percentage of disposable income. For instance, maximum 1/3 of disposable household income or maximum 70% of value of the underlying property.

- Regulation of hedge funds
  Hedge funds were not a primary cause of the crisis despite the enormous leverage they and the complexity of instruments they were using. On the contrary, they played a role in improving the liquidity of the markets over time. Measurements of systemic risk on the banking side would be sufficient to keep this industry contained.
- Sanctions of tax havens
  This is an irrelevant measure which will not solve future financial crises.
- Specific regulation to prevent systemic risk including the limitation of leverage
  The exponential rise of OTC derivatives contributed to the use of systemic risk. Certainly, the rapid growth of the credit derivative market should have been placed under control. The current ongoing discussions to set up a clearing house for credit derivatives would be a step in the right direction.

  The use of leverage should be left to hedge funds and investment banks. Commercial or retail banks should have limited ability to leverage up balance sheets. This could be arranged under the Basel Accords.

  Furthermore, Basel II should be stricter on the use of internal models to assess credit risk.
- Regulation to curb boom and bust cycles
  Trying to prevent boom and bust cycles will be as good as impossible, as this is inherent to our system. We have to go as far back as the 1630s in Holland during the Tulip mania to detect one of the first major bubbles mankind has known. Two hundred years later there was another major crisis, known as the Californian Gold Rush, which caused major disruptions in economic life (1840–50). Later on there was the Great Depression (1930s), and the Roaring Twenties which were the run up to that crisis. More recently, we have known the dot-com bubble of the late 1990s and the Great Credit Crisis 2007–8.

  As we laid out in this chapter, central banks play (unintentionally) a major role in creating bubbles. Abandoning the use of central banks, as Milton Friedman once suggested, would certainly not resolve the issue as we have known bubbles even before the existence of central banks.

  In order to be more pro-active in dealing with speculative bubbles, central banks should pay more attention to detect the build-up of a boom and bust cycle, and adjust monetary policy accordingly. In this respect they should pay much more attention to the development of asset prices in formulating their monetary policy.

  This will be a very challenging exercise.
- Stricter control of credit rating agencies
  More supervision of credit rating agencies is a necessity. Changing their fee driven business model, in which they could be held accountable for poor rating advice, would be a step forward.
- On an international level the G20 should convince Asian countries to use their build-up reserves to boost internal demand. The Asian economies were far too export driven via local currency manipulation, which in turn caused a global

trade imbalance. These excesses need to be worked away and should be prevented in future.

- A strict control on merger and acquisitions should be applied. In the future, regulators should avoid banks becoming 'too big to rescue'. A bank exceeding the GDP levels of its own country is a worrisome development and increases systemic risk. Smaller banks will also be to the benefit of the client as they are more relationship driven.

The Financial crisis of 2007–2008 demonstrated how banks had not learnt the lessons of earlier crises, such as the 1980–81 banking crash in the US. It remains important to learn well the lessons of the latest crash, so as to mitigate the downside risk in future.

Index

A
acceptance financing 30
accounts, corporate 430–3
  consolidated 433
administration 275
  and management of portfolios 275, 277, 280, 282, 384, 404
agencies, US 53, 62
Alpine bonds 55
American option 9, 344, 346, 348, 371–2
amortising loans 254
  negative amortisation 258
amortising structures 242
amortising swaps 326–17
Anglo Gold 442
annuity indexed bonds 206
annuity swaps 216–17
arbitrage
  models 164, 171–2
  opportunities 83, 111, 157
  pricing theory 84n17, 452
  proof 311–2
  transactions 273, 274, 276–8, 341
arbitrageurs 47
Asian financial crisis 58, 507
asset-backed bonds xxii, 237–8, 254, 275, 280
  see also collateralised debt obligations
asset-backed commercial paper (ABCP) 33, 35–9
asset and liability management 115, 251n6
asset-linked swaps 335–6
asset swaps 417–18
assets
  corporate 286
  fixed and floating 6, 444
  pool structuring 279
  prices of 353
  reference 275, 283, 285, 286, 405, 416–17
  sales in securitisation 237, 240, 241–3, 247
  types for CDOs 276–7
  see also collateral
at-the-money 9, 345, 346, 364, 370–1, 379
auction cycle 57n3
Australian government securities 50, 52, 205, 221
Austrian government securities 221
auto loans 36, 237, 243, 275, 276
Avellanada, M. 85

B
balance sheet 444
  capital management 239
  collateralised debt obligations 273–5
  corporate 430–2
  issues affecting 238, 245, 246, 273–5, 280, 283, 385
Bank of America 407, 419, 516
Bank of England 31, 56, 119, 120, 213, 249, 509, 522
Bank for International Settlements (BIS) rules 239, 284n9
Bank of New York 239
bank bills 25, 30
bank loans *see* loans
bankers acceptances 30–1
  eligible and ineligible 31
bankruptcy 30, 383, 387
  distancing from 35–6, 240
banks 30–1, 58–9, 115–16, 237–8, 273–5, 280, 295, 327, 384–5, 406, 486
  capital requirements 34
  central 48, 249
  clearing 27
  commercial 32, 34
  deposit accounts 509n18
  investment 34, 56, 58
  merchant 58
  mortgage 246
  OECD 286–91, 524n12
  retail 47–8
barrier credit option 404
Basel II 239, 524, 525, 526, 529, 530, 531
Basel Committee 475
Basel rules 239, 239n3, 284n7, 286
BAW model 362
benchmarks 31, 32, 50, 51, 187, 191, 262, 273
beta 441–2, 452–3
bid price 48
bid rate 26

bills of exchange 30
binomial tree 171, 479
Black, F. 382
Black 76 model 332, 382
Black–Derman–Toy model 167, 362
Black–Scholes model 8, 332, 351–5, 363–4, 367, 372, 374, 376, 378, 378, 413–14
Black–Scholes–Merton model 164, 410
Bloomberg 318, 338, 394, 397, 402, 403, 407
bond/bonds 46–120, 123–37, 139–48, 229, 231, 306–7
  asset backed xxii, 237–51, 254–71, 275, 280
  basis point value 128
  bullet 4, 299
  callable 55, 129, 259, 269, 270
  classes of 49, 75,
  corporate 14, 46, 53–5, 57, 63
  coupon 77–81
  default-free zero-coupon 23
  domestic 55
  duration 123–9
  Eurobonds 55–6, 59–60, 66, 67, 75
  foreign 56, 329
  global 47
  government 14, 46–7, 49–53, 56–7, 285, 347, 452
  inflation-indexed 203–11
  interest-only 260–1
  international 55–6
  issuance methods 50, 219–20
  mortgage-backed xxii, 86, 238–9, 254–72
  municipal 53
  options on 347, 359–60
  perpetual 67
  pricing and yield 63–85, 123–35, 172, 177, 206–14
  principal-only 260–1
  secured and unsecured 4, 55
  surety 37
  swaps 214
  synthetic index-linked 216
  undated 67
  underwriting of 58–9
  US agency 52–3
  zero-coupon 23–4, 66–7, 77–85, 100–4, 205, 301
Bond Market Association 262
bondholders, interests of 429–30, 430n1
bootstrapping 21, 89–96, 176, 183, 187–8, 218, 229–31
bought deal 59, 59n5
Brace–Gatarek–Musiela model 365
Brazilian government securities 221
breakeven formula 299
breakeven swap 216
Brennan, M. 159, 176
British Bankers Association 26, 297
brokers 49
  inter-dealer 49
Brown, Robert 173
Brownian motion 82, 160–1, 163, 169–70, 170–1, 173–4, 354, 380
B-spline methodology 188–201
building societies 27, 48, 115–16, 313, 476
Bulldog 55
bullet bond 4
business days 232

C

calibration of models 364, 409, 411, 413, 415
call aspect of mortgages 259
call option 8, 341–4, 355–7
  definition of 8, 341
  schedule 55
  see also options
Campbell, J. 88
Canadian government bonds 50, 51, 204n1, 205, 206, 221
capital
  controls 46
  cost of 3, 5, 435, 451–3, 452
  demand for 5
  gain 11, 68
  indexing 205
  requirements 31
  share capital 431–2
  sources of 46, 430
Capital Asset Pricing Model 362, 440–1, 451–3
capital buffer 245
capital market 46, 56
  structure of 3–9
caps 381–2
Carleton, W. 176
carry 68
carry trade 511–14, 513n23, 529, 530
cash
  current value of 11
  income 11
  instruments 9
  management 115
  reserve accounts 243
  waterfall 244–5, 275, 279, 280–1

cash flow analysis 14
cash flows 6
  of bonds 123–6, 139, 177, 204–6, 211
  of CDOs 278–82
  of credit-linked notes 391–2
  expected 36
  fair valuation of 23
  of interest-rate swaps 315, 316–7
  of mortgage-backed securities 261–6, 267–9
  of mortgages 255
  statement of 436
cash market instruments 6, 9
certificates of deposit 25, 27
CDS Big Bang 402
CDS points upfront 402–3
Chicago Board of Trade 7, 8
Chicago Board Options Exchange 8, 347
Chilean government securities 221
clean deposits 26
clean price 65, 73–7
clearing house 8
collar 382
collateral 31, 33, 57
  for CDOs 276, 280–1, 284
collateralised bond obligations 273
collateralised debt obligations xxii, 278–81
  asset types 276–7
  cash flow CDO structures 278–82
  investor analysis 277–82
  market size 273–4, 275–6, 280, 283
  risk and return 287–9
  synthetic CDOs 275, 283–91
collateralised funding 397
collateralised loan obligations 273–4
collateralised mortgage obligations 261
Colombian government securities 221
commercial bills 30
commercial paper 25, 31–43, 273
  asset-backed 33, 35–43
  credit-supported 33
  issuance methods 34
  issuers of 33
  market size 39, 40
  purchasers of 33
  US and Euro compared 33
commissions 31
commodity markets 7
Companies Act 1985 (UK) 430
compliance tests 282
composite securities 6
compounding 15–17, 299
  continuous 16
concentration limits 485
conduits 35
consumer price index 206
continuous-time pricing 81–6
contract for difference (CFD) 221
contracts
  forward vs option 7 (*see also* forward contracts)
  standardised and tailor-made 7–8
convexity 129–36, 295
Cooper, I. 176
corporate entities (firms)
  as bond issuers 46, 54–5
  definition of 3
  financial structure 429–30
  financing arms of 33
  hedging by 306–7
  memorandum and articles of association 429
  money market instruments used by 25, 31–43, 336
  users of swaps 335–7
  valuation of 450–1 (*see also* equity)
correlations
  in reference portfolio 287
  in VAR 457, 459, 460–3, 482
cost of funding 238
Cost of Funds Index (COFI) 258
coupon 27, 65–7, 70–2, 80, 123, 348, 388
  bonds 70, 77–84, 86, 105–7, 205
  CD 27
  dates 66, 70–2, 73
  fixed 4
  payment intervals 70–1, 86n20, 204
  payments 69, 74–5, 77, 81, 211, 221
  process 82
  zero *see* zero-coupon bonds
  see also bond pricing
Cox, J. 109, 112, 116, 166, 176, 310,
Cox–Ingersoll–Ross model 166–7, 176, 310, 311–2, 365
credit
  arbitrage vehicles 36
  catastrophic outcomes 478
  curves 421–2
  default swaps 283–95, 384, 385, 386–9, 421, 422–35
  derivatives *see* derivatives, credit
  enhancement of 36–7, 240, 243–4, 275, 277, 281
  event 383, 386–7, 391–2, 392
  intermediation 384
  limit setting 484
  linked note 386, 391–3

credit (*continued*)
migration 411
portfolio diversification 384
premium 52, 103
quality 27
ratings 27, 31, 62, 240–1, 248, 255, 275, 277, 278–9, 285, 286, 384, 388, 483, 485
returns vs market returns 479
risk/exposure *see* risk, credit
spread *see* spreads
stand-by lines of 34
support 33
switches 405
credit card loans 36, 237, 243, 275, 391
Credit Suisse First Boston (CSFB) 479
CreditManager 479, 483–84
CreditMetrics 477–84, 587
CreditRisk+ 479
creditors, secured and unsecured 3–4
cross-rate 39–40, 41
forward 44
cubic B-spline methdology 188–201
cubic polynomial, use of 178–9
Culbertson, J. M. 115
currency issues 6, 327–8, 406
see also foreign exchange
currency swaps 328–31
Czech Republic government securities 221

D

Danish government securities 221
Das–Tufano model 411–2
day-count conventions 26, 27, 40, 49, 75–6, 186–7, 320n1, 232–3
dealer paper 34
dealing arrangements 7–8, 296, 316
pit trading 7–8
see also over-the-counter, exchanges
debt finance
classes of 275
definition of 4
effect of gearing 447–8
maturity of 4
original features of 6
debt management policy 120
debt service coverage ratio (DSCR) 248
default-free investments 14
default fund 8
defaults
actual 384, 387, 479
on CDOs 287
credit risk 477
on loans 9, 14, 60–3, 259, 383–4, 392
probability rates 287, 410, 412–3
Standard Default Assumption (SDA) 265
technical 383–4, 387
deficits, government 46
deliverable obligation 386, 389
delta 372–3, 377
demand curve 5
deposits, money market 26–7
derivative instruments 275, 283
advantages of 9
applications 219–20
credit *see* derivatives, credit
definition of 3, 7
financial 7
inflation-indexed 213–20
interest rate 172, 295–312
main types of 6
specialist 7
derivatives, credit 383–25
classification of 386–7
credit default swaps *see* swaps, credit default
credit-linked notes *see* credit-linked note
credit options 416–7
pricing and valuation 409–25
total return swaps *see* swaps, total return
uses of 384–5, 389–90, 395, 404–8
Deutsche Terminbourse 303
digital credit derivative 388
dirty price 66, 74–7, 86, 417
discount
basis for money market instruments 26, 29–31
curve 85
equation 103
factors 19–21, 84, 101, 102, 176, 186, 187, 199, 325
function 19, 86, 100–5, 177
instruments 34
rate 11, 64, 435
discounting frequency 72
disintermediation 385
disruption, market 37
divergence 435
diversification
of funding sources 60
of portfolios 384, 441
diversity tests/score 282, 287
dividends 6, 11, 348, 350, 363, 364, 429
growth model 435
optionality of 6
policy on 437–8
valuation model 435

dot-com market 430, 437n4
due diligence 242, 247
Duffie–Singleton model 412–3
duration 123–9
  of mortgage-backed securities 267–71
Dutch auction 49

E
earnings, economic 436
econometric techniques 105
efficient markets hypothesis 440–1
endowment mortgages 254
epsilon 376
equilibrium 109
  concept of 4
  models 171–2
  rate of interest 11–12
equity 3–4, 363–4, 449–75
  classes of shares 429–30
  default swap xxii
  distinction from debt 3–4, 6, 429
  essential features 6, 429
  piece/note 243
  preference shares 4, 6n4
  premium 431453
  repayment of 4
  return on 446–7
  valuation of 434–7, 440–3
eta 376
EURIBOR 26
euro, UK possibility of joining 119
euro bond market 57–8
Eurobonds 55–6, 59–60, 66, 67, 71, 75
Eurodollar deposits 46
European Central Bank 26
European market 239, 286, 387–8
European option 9, 332, 344, 363, 371–2
European Union 55, 251
ex dividend 74
excess spread 244
exchange delivery settlement price (EDSP) 304
exchanges 346–7
  see also futures exchanges, stock exchanges
exercise price 8
expectations hypothesis 107–12
expected earnings valuation model 436–7
expenses, corporate 432–3
exposure limits 484–5

F
Fabozzi, F. 262, 264
Fama, E. 5, 353, 440
Farm Credit System 53
federal funds (fed funds) 32, 508
Federal Reserve 56, 63, 249, 493, 496, 505, 507, 508, 520
Federal Home Loan Bank System (FHLBS) 52
Federal Home Loan Mortgage Corporation (Freddie Mac) 52–3, 255, 520–2, 525, 526
Federal National Mortgage Association (Fannie Mae) 52, 53, 255, 520–1, 525, 526
Federal Reserve Bank (Fed) 31, 32, 63, 508–9
finance, corporate sources of 430
financial companies 33
financial derivatives *see* derivative instruments
financial engineering 47, 254
Financial Services Authority (FSA) 56
Financial Times 449
firm *see* corporate entities
first-loss piece 243, 285
Fisher, I. 5, 12, 108, 440
Fisher equation 12
Fisher identity 212–13, 204
Fisher–Hirschleifer model 440
Fitch 248, 277n2
fixed income securities 46–158
fixed interest securities 6
fixed price re-offer mechanism 58–9, 58n4
fixing date 298, 299
floating-rate notes 86
floors 381–2
flow trading 390
Fong, H. G. 160, 188
foreign exchange 39–5,
  forwards in 7, 41–5
  options 347
forward contracts 7, 91, 301–2, 361
  credit spread 415–16
  for foreign exchange 41–5
  non-deliverable 42
forward-forward swap 44–5
forward rate 42, 104, 295, 320–23
  agreement (FRA), 295–301, 303 472–3
  implied 90–1, 304–305, 320
forward-start swap 344
forward swap 43–5
France futures trading 303
free market economy 4
French government bonds 50, 51, 52, 79, 221

frequency of interest payments 15–16
FTSE-100 347
fund management 47, 115, 211, *see also* portfolio management
funded instruments, 403
funding
  cost of 241
  instruments 396–7
  mix, diversification of 238–9
funds, flow of 3
funds, use of 436
future values 15
futures 7–8, 302, 342, 347, 361
  basic trading model 7
  exchanges 5, 7–8, 302, 303
  options 347,
  prices and actual cash rates 305
  settlement of contracts 310
  short-term interest rate 302–310
  standardisation of contracts 7–8

G
Galitz, L. 9
gamma 373–5
Garman and Kohlhagen model 362
Gaussian diffusion process, 365
gearing 378,
Germany
  futures trading 303
  government bonds 50, 51
gilts 13 64, 66, 67, 68, 90, 150 119–20, 191–2, 204n1, 205, 221, 363
Global Master Repurchase Agreement (GMRA) 395
globalization 56
GMRES 193–9
gold market 442
Gordon growth model 435
Government National Mortgage Association (GNMA) 53, 255
governments
  economic policy 119–20
  as raisers of finance 3, 46–7, 59
  see also bonds, government
Grabbe model 362
Greek government securities 221
Greeks, the 372–80
Greenspan doctrine 508
Greenspan put 508, 530
gross redemption yield *see* yield to maturity
guaranteed return 6
guarantees, bank 284
guarantors 248

H
Harrison, J. 353
Harrison and Pliska model 365
hazard rate 412–13
Heath, Jarrow and Morton model 167–70, 365
hedge funds 486
hedge, long and short 306
hedge ratio 309–10, 374–5
hedging 9, 44, 110, 128, 129, 168, 172 203, 214, 220, 286 295, 299 313, 341–82,
  bond instruments using swaps 337–40
  delta hedging 373, 377
  of interest-rate futures 304–310, 331
  of mortgage-backed securities 271
  of pension liabilities 219–20
  using credit derivatives 383
Hicks, J. 108
Hirshleifer, J. 440
historic cost basis (corporate finance) 431
historical simulation 459, 467
Ho, T. 161
Ho and Lee model 164–5
Hull, J. 161
Hull–White model 162, 165, 167, 363
Hungarian government securities 221
hybrid securities 6

I
Iceland government securities 221
importing and exporting 30, 47
in-the-money 9, 331, 345, 352, 370–1
income, overall levels of 5
income statement, corporate 432–3
Index credit default swap 406
index-linking 203–23
  choice of index 203, 206
  lags 204, 208, 212, 217
  mortgage interest 258
  options 347
  rebasing 207
Indian government bonds 67
indices, stock market 430, 452
inflation
  break-even rate 210
  expected rate of 5, 111, 211, 212–13
  HICP measure 215
  impact on interest rates 12–13
  indexation 203–223
  indicators of 5
  term structure of rates 213, 217–19
information technology, impact of 46–7

Ingersoll, J. 64, 94, 108, 109, 112, 116, 166–7, 176, 310, 311, 355, 365
insurance
  companies 48, 115, 244
  life 48, 115
  options as an 341
  pool 244
inter-dealer brokers (IDBs) 49
interbank market 17
interest 11
  accrued 66, 74–7, 80n13
  on CDs 27
  compounding of 15–17, 83
  coverage 282
  discount rate 65
  gain 11
  indexing 204–5
  levels of payment 4
  on mortgages 254, 255–6
  nominal rate 12
  prime rate 32
  reinvestment of 15
  simple 14
interest rate 352, 378
  actual market 71
  annualised 14–15, 16, 17, 64
  and bond prices 108, 109, 111, 119, 123–38
  calculations in the money markets 26
  changes in levels 5
  changes and refinancing 55
  constant assumed 363
  derivatives 172, 295–312
  fixed 6, 8
  floating 8
  forward 71, 310–11
  implicit on discount instruments 29
  implied 71
  instantaneous 82
  market-determined 11–14, 73–4
  modelling 159–74
  negative 166, 166n7, 167
  options 370, 381–2
  real 211–3
  risk 123–38, 286
  risk, second-order 133
  risk-free 348, 355
  short(-term) 82, 84, 159, 164, 168,
  short-term derivatives 172, 295–312
  short-term futures 302–10
  swaps 8, 313–40
  real 5, 13
  variety of 12n1
  volatility 160–1
  zero-coupon 71, 186–7
intermediaries, financial 47–9, 59
internal rate of return 68, 69, 90
International Petroleum Exchange 8, 302
International Swaps and Derivatives Association (ISDA) standard 315, 387, 387n1, 397, 402, 424
interpolation 18–19, 101, 104, 177–8, 181, 199
investment, corporate 436–7
investment funds 25
investment grade rating 36, 63, 241
investors
  basic forms of compensation for 11
  in collateralised debt obligations 277–82
  expectations of 108–9
  individuals as 25, 203
  institutional 48
  in the money market 25
  preferences/requirements of 54, 71
Irish government securities 221
ISDA *see* International Swaps and Derivatives Association
ISO convention 39
Italian government securities 221
iterative methods 70, 194–5, 213
Ito's lemma 162–163
iTraxx index 398, 399, 400, 401, 497, 497n4, 499

J

James, J. 105, 168, 184
Japanese
  bank CLOs 274
  market 49, 55, 58, 166
  yen market 58
Jarrow, R. 87, 88, 109
Jarrow, Lando and Turnbull model 411
Jensen's inequality 109
Jordan, J. 183
JPMorgan 36, 403, 477, 496, 515

K

kappa 372
Kessel, R. A. 114
Kreps, D. 353

L

lambda 378, 379
Langetieg, T. C. 177
leases, real-estate 36
leases, receivables on 36, 237, 273

Lee, S.-B. 161
lender of last resort (LoLR) 501, 502, 509
leptokurtosis 463, 487–8
letters of credit 33, 37, 500n6
leverage 69, 378, 386, 446–7
liabilities, corporate 432, 433
liability-linked swaps 336
Libid 26
Libor *see* London Interbank Offer Rate
life assurance companies 48, 115
LIFFE 8, 300, 302–3, 347
limited liability 60
linear relationship 78n12
liquidity 6, 21, 27, 34, 52, 177, 203, 200, 304, 301, 311, 410
  preference theory 112–4
  premium 13, 113, 117
  risk 291, 502
  rules regarding 52
  support 37
Lisbor 26
Lo, A. 88
loans
  auto 36, 237, 242, 254, 275, 276
  bank 4, 31, 35, 239, 283, 385, 406–7
  collateralised obligations 273
  consumer 242
  corporate 36
  irrevocable facilities 37
  mortgage 36
  servicing 276
  underwriting 242
local authorities 3, 25, 27
  as bond issuers 49, 53
lognormal distribution 363, 366–7, 467
London Business School 441
London Interbank Offer Rate (Libor) 26, 36
London Metal Exchange 302
Longstaff and Schwartz model 180
Long Term Capital Management 486, 511, 517

M

Macaulay's duration 123–7, 267
macroeconomics 1, 77, 277
managed funds 47, 115, 211, 274
management, corporate 438
management fees 287
Mansi, S. 183
mapping of cash flows 464–5
margin
  definition of 8
  swap 327
  transfer 347
market(s) 15
  bond 56–60, 210
  commodity 7
  complete 353
  efficiency scenarios 353, 440
  European 283, 286
  extreme events 475, 476
  financial, make-up of 5–6
  frictionless 118, 363
  grey 59, 59n5
  imperfections 172
  instruments 6, 25–45
  issuers 25
  making 47
  mechanism for financing price 4–5
  money *see* money market
  over-the-counter *see* over-the-counter
  perfect 118
  prices 179
  primary 46, 57
  professionals 48–9
  repo 56
  returns vs credit returns 478
  secondary 27–30, 33, 57, 59, 66, 314
  sectors, gaining exposure to 405
  sizes of different 46, 49, 145
  UK 15, 57–8, 214, 297, 302, 346
  US 15, 57–8, 278n3, 283, 347
  value basis 431
  when-issued 56, 57
  wholesale and retail 17
market-to-book (MB) ratio 448–50
marking to market 323, 477, 484
Markowitz, H. 441
martingale method 353
master trusts 243, 274n1
Matador bonds 55
MATIF 303
matrices
  for GMRES 193–7
  transition for pricing derivatives 411
  for VaR calculation 460–1, 472
maturity 25
  band for collateral 31
  date 297, 300
  definition of 4
  impact of period 16, 114, 213
  mismatches 388, 393
  in money and capital market 56
  for mortgage-backed bonds 238
  and option value 348, 349
  redemption at 54, 55
  and registration requirements 32
  standardised 18

time to 85, 111–12
yield to 68–74, 86, 207
MB ratio 450–1
McCulloch, J. 106, 177
McKinlay, A. 88
Merton, R. 164
Merton models 410
Merton's bound 357
Merton, Barone-Adesi and Whaley model 362
Mexican government securities 221
Microsoft
as corporation 6, 437
products of 325, 335n4, 338, 359, 367–8, 462n1
Miller, M. 437, 438
minimisation problem 190–1
Minsky moment 495
models 409–13
B-spline 186–201
BAW 362
binomial 269–71
Black 76 332, 360–2, 382
Black–Derman–Toy 167, 362, 365
Black–Scholes 8, 332, 351–60, 363–4, 367, 372, 373, 376, 378, 378–80
Black–Scholes–Merton 409–10
Brace–Gatarek–Musiela 365
calibration 176, 364, 365, 411
cash flow 222
choice of interest-rate 170–3
Cox–Ingersoll–Ross 166–7, 184, 365
of credit risk 493–502
credit spread 413–5
Das–Tufano 411–2
Duffie–Singleton 412–3
equilibrium 171–2
financial/cash flow 222, 267–9
Fisher–Hirschleifer 440
Garman and Kohlhagen 362
Gordon growth 435
Grabbe 362
Harrison and Pliska 365
Heath, Jarrow and Morton 167–71, 365
Ho and Lee 164–5
Hull–White 165, 167, 363
interest rate 159–74
Jarrow, Lando and Turnbull 410–1
Longstaff and Schwarz 180
mean reversion 414, 416
Merton 410
Merton, Barone-Adesi and Whaley 362
multi-factor 365
prepayment 255
reduced form 410
risk premium 452–3
Roll, Geske and Whaley 362
for share valuation 450–3
simulation 287, 467–9
stochastic 159–74, 411, 487
stress testing 475–6
structural 409–10
term structure 159–74
of VaR 461–81
Vasicek 161, 163–4, 365
Waggoner 213
Modigliani, F. 116, 437, 438
monetary policy 118
money market
availability of funds in 118
composition of sterling 40
maturity in 56
rates 100, 104, 166
money market instruments 25–45
discount-based 29–31, 34–5
types of instrument 25–26
yield-based 26–8
money substitute hypothesis 114–15
Monte Carlo simulation 171, 459, 467–9, 475, 487
Moody's 63, 248, 287, 387, 407, 479, 484
mortgage-backed securities xxii, 86n20, 238, 242, 254–71, 278
cash flow patterns 261–6
evaluation and analysis 266–71
types 260–1
mortgages 237, 254–71
balloon 258
commercial 260
endowment 254
fixed and floating rate 254
graduated payment (GPM) 258
growing equity (GEM) 258
hybrid 258
repayment and interest-only 254
risk 259
servicing fee 256–7
multi-tranching 407
mutual funds 33

N

National Cost of Funds Index 258
NationsBank 274
Neftci, S. 84
negative convexity 270
negative funding/negative carry 68
negotiability 27
Nelson, C. 105

Nelson and Siegel curve 183–4
New York Stock Exchange (NYSE) 52, 347
New Zealand government securities 205, 205, 221
no-arbitrage environment/condition 23, 42, 88, 89, 109, 163, 168, 301, 302, 353, 410
non-parametric curve fitting 180–4
normality, statistical assumption 458, 487–8
Northern Rock plc 251n
Norwegian government securities 221
note classes 244, 279–80, 284
  first-loss piece 243, 284
  junior 244, 285
  mezzanine 285
  senior 244, 279
note, credit-linked 391–2

O

off-balance sheet instruments 9, 104, 128, 295, 302, 314, 341, 375, 385, 395
off-the-run securities 52
offer price 48
offer, private 56
offer, public 56
oil, price of 7, 46
Oldrich, A. 188
on-the-run securities 52, 57
open-market operations 120, 120n36
option-adjusted spread analysis 269–70
options 8–9, 341–82
  at-the-money 9, 370, 371, 374, 375, 377, 379
  Bermudan 344
  call and put 8, 53, 55, 219, 331, 341, 343–4, 345, 382
  credit 411, 416–7
  currency 362
  definition 346–7
  European and American 8, 344, 360, 363, 371–2
  futures 347
  hedging 371
  in-the-money 9, 331, 346, 354, 370–1, 379
  instruments 346–8
  on interest rates 347, 381–2
  out-of-the-money 9, 331, 346, 354, 370–1, 379
  physical and financial 344
  premium 345
  price behaviour 370–2
  pricing 348–50, 351–3
  put–call parity relationship 357–8
  risk measurement 372–8
  smile 378–80
  swaptions 326–7, 331–5
  terminology 344–6
  value 345, 370–2
  variables 372–8
ordinary least squares technique 106, 177
Ornstein–Uhlenbeck process 162
out-of-the-money 9, 331, 346, 354, 370–1, 379
over-collateralisation 37, 242, 281, 282
over-the-counter (OTC) trading/products 7, 52, 52n1, 57, 217, 295, 301, 313, 346, 347, 383
overdraft, bank 4
overnight rate 18

P

par value 6, 54, 417–8, 429
parastatals 3
pass-through structures 242
pay-as-you-go swap 216
payoff profile
  linked to another security 6
  of options 8, 342–4
peer group 450–1
pension funds 33, 48, 54, 115, 203, 216, 219, 337
Philadelphia Currency Options Exchange 347
pip 40–1
pit trading 7–8
Pliska, S. 353
Polish government securities 221
pool insurance 244
portfolio
  approach to credit risk 477
  concept for VaR 455
  diversification 384, 441
  effects 477
  management/administration 276, 278, 279, 385, 390, 404–5
  restructuring 220
  sample bond for calculating VAR 469–72
  theory 440–1
portfolio credit default swap 390
pre-payments 242
preference shares 4, 6n4
preferred habitat theory 116
prepayment risk 251, 259–60
present value 14–19, 24, 187
  of a basis point (PVBP) 337–40
  rule 440
price–earnings ratio 448–50
price premium 417

pricing 3, 172, 299–301, 304–5, 309, 332–5, 348–9, 351–68, 380, 385, 409–23
  bonds 64–86, 105–7, 123–37, 172, 177, 269–71
  commercial paper 39
  equities 434–7, 440–3
  futures 303–5, 311–2
  market approach 37–9
  of options 348–50, 351–3
  of swaptions 334–5
  of Treasury bills 29–30
  of zero-coupon swaps 318–26
  see also yield
principal–agent issues 438
private securities 278, 278n3
probability theory 160
profit
  corporate 6, 429–30, 444
  and loss account 432–3
  of proprietary traders 48
  ratios 444
  retained 431
proprietary traders 48
public sector bodies 3, 47, 53, 56–7, 59
  bond issues by 56–7, 63–4
Public Securities Association (PSA) 262, 265
put options 8, 53, 326, 341, 344, 346

Q
QuantLib, 96, 224–33
quanto option 328
quote spread 26

R
ramp-up period 277, 286
random walk 441, 467
RATE application software 382
rate of return, overall 11, 71
rating of investments 240, 248–9
rationality, assumption of 440–1
ratios
  analysis of 440–53
  management-level analysis 446–50
  OBS instruments to improve 35, 238
receivables, trade/equipment lease 36
recessions 118–9, 119n34
recovery rate 62, 287–8, 410–3, 479
redemption
  date 67
  price 54–5
  value 6
  yield 68
reference assets *see* assets
reference rate 297, 326
refinancing 55
regression techniques 106, 176–7
regulatory environment/issues 5, 31, 56, 59n7, 115, 239, 284, 286
reinvestment of income 15, 71, 73
rents 11
repo *see* repurchase agreements
repo market/rate 56, 214
repurchase agreements 25, 166, 395
  synthetic repo 395–6
  repo trade 396
reserves, corporate 432
residual claim of shareholders 430
Resolution Funding Corporation (REFCO) 53
restructuring 387
Retail Price Index (RPI) 203, 206
retained earnings 4
return on capital 3, 4–5, 443–4
return on equity 446–7
return on net assets 444, 446, 447
return, real 5
Reuters 39, 296, 316, 318, 394
revolving structures 242–3
rho 378
risk
  basis 327
  counterparty 8
  credit xxii, 9, 11, 25, 34, 53, 60–3, 240, 283, 283–7, 286–7, 384–5, 395, 404–5
  credit default 477
  credit spread 477
  credit and return equation 4
  default 14, 60–1, 104
  with equities 429, 435, 441
  exposure 240
  foreign exchange 60
  integration of credit and market 486–7
  interest rate 123–38, 268, 286
  investment appraisal 441
  liquidity 487
  management 240, 457–88
  market 452
  measurement 457–88
  mitigation 484
  mortgage 259
  portfolio 441, 477
  premium 14, 104, 441, 453
  reinvestment 205
  residual 475
  reversal 371
  security 9
  of sinking funds 54
  VaR methodology for credit 486–7

risk-adjusted return on capital (RAROC) 479–80
risk-free bonds, theoretical 24
risk-free investments 64, 412, 415–6
risk-free yield 29
risk-neutral measure 84, 353, 410
risk/reward profile 295
  of options 342–3
RiskMetrics 457, 462, 463, 464, 465, 470, 472, 473, 474, 487
Roll, Geske and Whaley model 362
rollover 34, 35
Roodepoort Deep 442
Ross, S. M. 84, 109, 112, 116, 160, 166–7, 176, 311, 437, 443
Rozeff, M. 460
Rubinstein, M. 110
running cost 68

S

sales ratios 444
Samurai bonds 55
savings
  overall level of 5
  total supply of 5
scenario analysis 478
Schwartz, E. 159, 176
Securities and Exchange Commission (SEC) 32, 56
securities, definition of 3, 6, 9
securitisation 4, 6, 35, 237–53, 273–4, 283, 406
  advantages 35
  definition 238
  market growth 238
  note tranching 243
  process 240–9
  structures 247–9
  synthetic 237
segmentation hypothesis 115–6
self-liquidating transactions 31
setting date 316, 327
settlement
  cash or physical 388–9, 424
  date 66, 75, 297
  mechanism 386, 424
  sum 297
shadow banking 491, 494, 500–2, 513, 526, 529, 530
shareholders, interests of 429–30
shareholders' funds 431
shares *see* equity
Sharpe, W. 441
Shiller, R. 108
short selling 385, 395n5
Siegel, A. 183
signalling, corporate 437–8
SIMEX 347
Singapore
  government bonds 50, 52
  options trading 347–8
sinking fund 54
size of
  bond issues 58
  funding requirements 4
South Africa
  gold market 442
  government bonds 50, 52, 67
Sovereign wealth funds, 491, 502–17
special purpose corporation (SPC) *see* special purpose vehicle
special purpose vehicle (SPV) 35, 36, 237, 239, 240–53, 275, 278, 283, 284
structures for 242–3
speculation 9, 47, 295–6, 341
spline estimation methods 106–7, 178, 180–3, 186–201
spot date 298
spot rates/trading 18, 26, 44, 86, 88, 105–7, 319, 361
for foreign exchange 40
spreads 40–1, 48–9, 52, 62, 177, 318, 394, 409–10
  credit spread models 413–15
  credit spread products 415–18
  credit spread risk 477
  option-adjusted analysis 269–70
  trading in credit 406
Standard & Poor's 63, 248
  S&P500 347, 452
Standard Chartered Bank 476
Standard Default Assumption (SDA) 265
statistical concepts, basic summary 366, 458, 487–8
Sterling bond market 55
Stibor 26
stochastic
  models 159–74, 411, 487
  rates 84
  volatility 363
stock exchanges 5, 52, 449
stress testing 475–6
strike price 8, 344, 346, 348, 352
stripping 206
strips 67, 77, 306
  see also zero–coupon bonds
structured financial products xxii, 36, 235–92

structured investment vehicle (SIV) 37, 498, 502
Student Loan Marketing System (Sallie Mae) 53
sub-prime mortgage
delinquency, 494
Sundaresan, S. 78, 237
supply curve 5
supply and demand, laws of 3, 4–5
supranational bodies 3, 59
Sutch, R. 116
Svensson model 183
swap/swaps 313–40
applications 331–40
asset 417–8
asset linked swaps 335, 337
breakeven 216
credit default 281–90, 388–91, 411
currency 328–31
definition of 8
forward 43–5
forward-forward 44–5
forward-start 328
as hedge tools 337–40
index 406
inflation-linked bond 214–5
interest-rate 8, 313–28, 331, 337–40, 473–5
liability-linked 336
long and short sides 314–5
margin 327
non-vanilla 326–8
pay-as-you-go 216
points 43–4
rates 104
spread 62, 318
term sheet, sample 423–5
termination 394
TIPS 216, 221–3
total return 392–6, 405, 411, 418
types of 313, 326–8
valuation 323, 328–9, 332–5
year-on-year inflation 215–6
zero-coupon 216, 318–26
swaptions 172, 326, 331–5, 360
Swedish government securities 205, 221
SWIFT 39
syndicates, bank 58
synthetic instruments 215, 275, 283–91
synthetic repo 395–6

T
tailing 309
Taiwan government bonds 50, 51
tax clientele effect 438
tax-exempt securities 53
taxation 34, 59, 63, 205, 438
corporate 5
liability for 6
personal 5
Taylor expansion 137–8
TED spread, 495, 496
Tennessee Valley Authority (TVA) 53
term
structure of inflation rates 213, 217–9
structure of interest rates 78, 88–97, 310–1, 365
structure modelling 159–74
typical for CP 31, 33–4
typical for corporate bonds 53–4
see also maturity
term sheet, sample 423–5
theta 375–6
tick 67, 347
time
decay of an option 375–6
deposits 26
factor in modelling 173
as function of credit survival probability 421–2
horizon for modelling credit risk 479
preferences 11–12
value of money 14–21, 68, 440
value of an option 370–1
TIPS/TIPS swap 216, 221–3
tom next 18
tractability 163n3
trade bills 30
trade date 297, 315
trading *see* dealing arrangements
transaction costs 311, 354
transition matrices 410–11
treasuries, corporate 34
Treasury securities 4, 25, 29–30, 63
see also bonds, government; US Treasury securities
trustee 54
Turkish government securities 205, 221

U
underwriting
process 58–9
for securitisation 241–2
unfunded instruments 386
United Kingdom
government bonds 14, 49–52, 67, 68, 71, 73, 79, 119–20, 191–3, 205, 206n2, 221
market 297, 302, 346–7
mortgage system 254, 258

United States 5
agencies 52–3, 63, 255
agency bonds 52–3, 63
government bonds *see* US Treasury securities
markets 15, 58, 255, 278n3, 283, 346–7
mortgage system 254, 255, 257–8
trade deficit 46
US Treasury securities 4, 32, 49–52, 67, 75, 79, 117, 205, 216, 221, 221–3, 347

V
value-at-risk (VaR) xxii, 457–88
applications of credit VaR 484–6
calculation methods 460–75
confidence intervals 465–66
for credit risk 477
value basis 305
value date 66–7, 74
vanilla products 4, 8, 325
bonds 86n20
variables
random 160
state 81–2, 168
variance-covariance 459, 461–6, 470, 471
Vasicek, O. 160, 161, 188
Vasicek model 163–4, 165–6, 176, 180, 365
vega 376–8
volatility, credit spread 416–17
volatility, market 54, 352, 363, 376, 421, 441, 452, 474
changes over time 487
implied 364, 379, 460
smile 363, 378–80
stochastic 363
volatility (as modified duration) 127
volatility, portfolio 458, 460
voting rights 6

W
Waggoner model 213
warehouse facility 286
waterfall 244–5, 275, 280, 281
Webber, N. 100, 104, 184
weighted-average lifer (WAL) approach 242
weighted average spread/rating 282
Weiner, N. 173
Weiner process 82, 160, 163, 168–9, 174
White, A. 161–2
winding up, implications of 6
see also bankruptcy, defaults
World Bank 3
writer (of options) 9

Y
Yankee bonds 55
year-on-year inflation swap 215–6
yield 18
annualised 73
of bonds 63–86, 210–3
basis for money market instruments 26–8
on CDs 27–8
of CP 34–5
current 68, 81
definition of 12
expected 73
flat 68, 81
gross and net 70
interest 68
to maturity 68–74, 69n9, 80–1, 207
measures for index-linked securities 207–8
of mortgage-backed bonds 267–9
real 211
risk-free 29
running 81
spreads, typical 62–3
yield curve 13–14, 88, 112–7, 305, 318, 476
analysis 107–20
choice of 184
downward-sloping 107, 118–9
expectations hypothesis 107–12
fitting techniques 159–74, 176–84
flat 49, 107, 113, 117–18
for government bonds 50, 52, 313
forward rate 88
humped 107, 111, 113, 183
inverted 79, 113, 117, 119–21
modelling 159–74
shapes 79, 107, 116, 118–20, 183
smoothing 176–80
for swaps 313, 318–19
upward sloping 13, 49, 107, 117, 119n35
zero-coupon 78–80, 176, 364–5
YieldCurve com xxv, xxvi, 170
YieldCurve.publishing xxv, xxvi

Z
zero-coupon
bonds 23–4, 66–7, 77–85, 100–3, 204–5, 301
interest rate 71, 186–7
rate 81, 102, 105, 177, 183, 186–7, 319–25
swap 216, 318–26