The Valuation of Interest Rate Derivative Securities

Jeroen F.J. de Munnik

ROUTLEDGE

THE VALUATION OF INTEREST RATE DERIVATIVE SECURITIES

The increased volatility of interest rates during recent years and the corresponding introduction of a variety of interest rate derivative securities like bond options, futures and embedded options in mortgages, stress the need for a comprehensive financial theory to determine values of fixed income instruments and derivative securities consistently.

This book provides:

- A detailed overview and classification of the different approaches to value interest rate dependent securities
- A comparison of the numerical approaches to value complex securities
- An empirical examination for the Dutch Fixed Income Market of some well-known interest rate models which demonstrates recent improvements to describe interest rate movements in relation to contingent claim valuation

Jeroen F.J.de Munnik completed his Ph.D. thesis The Valuation of Interest Rate Derivative Securities' in 1992. During his Ph.D. study he worked for J.P.Morgan Securities Inc. in New York and published in the *Journal of Banking and Finance*. Currently, he is working as a financial controller at Spaarbeleg Bank, a 100 per cent subsidiary of AEGON Insurance.

ROUTLEDGE NEW ADVANCES IN ECONOMICS

Series Editor: Albert Jolink

1. THE VALUATION OF INTEREST RATE DERIVATIVE SECURITIES
Jeroen F.J.de Munnik

THE VALUATION OF INTEREST RATE DERIVATIVE SECURITIES

Jeroen F.J.de Munnik

London and New York

First published 1996 by Routledge 11 New Fetter Lane, London EC4P 4EE
This edition published in the Taylor & Francis e-Library, 2005.

"To purchase your own copy of this or any of Taylor & Francis
or Routledge's collection of thousands of eBooks please go to
http://www.ebookstore.tandf.co.uk/."

Simultaneously published in the USA and Canada by Routledge 29 West 35th Street, New York, NY 10001

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

Library of Congress Cataloguing in Publication Data Munnik, J.F.J.de (Jeroen F.J.), 1966– The valuation of interest rate derivative securities/Jeroen F.J. de Munnik. p. cm.—(Routledge new advances in economics; 1) Includes bibliographical references and index. 1. Securities–Valuation–Mathematical models. 2. Interest rates– Mathematical models. 3. Derivative securities–Valuation–Mathematical models. I. Title. II. Series. HG4515.2.M86 1996 96–6089 332.63′2042–dc20 CIP

ISBN 0-203-98277-0 Master e-book ISBN

ISBN 0-415-13727-6 (Print Edition)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

Ω	Sample space
$\mathcal{F}$	σ-algebra
P	Probability measure
Q	Risk-neutral probability measure
F	Filtration
$(\Omega, \mathcal{F}, F, P)$	Filtered probability space
$Ep(. \mid \mathcal{F}_s)$	Expectation operator given the information at time s with respect to the probability measure P
$E_Q(. \mid \mathcal{F}_s)$	Expectation operator given the information at time s with respect to the risk-neutral probability measure Q
t	Calendar time
τ	Time-to-Maturity
T	Final Trading Date
$r(t)$	Instantaneous short-term rate of interest at time t
θ	Unconditional mean of instantaneous short rate
κ	Level of mean reversion of instantaneous short rate
σ	Volatility of instantaneous short rate
λ	Market price of risk
$\rho(\lambda, t)$	Radon-Nikodym derivative at time t
$W(t)$	Standard Brownian Motion at time t relative to the probability measure P
$\tilde{W}(t)$	Standard Brownian Motion at time t relative to the risk-neutral probability measure Q
$f(t, \tau)$	Instantaneous forward rate with time-to-maturity τ at time t
$R(t, \tau)$	Spot rate with time-to-maturity τ at time t
$\mu_R(s, t)$	Mean of spot rate with time-to-maturity τ at time t given the information at time s
$\sigma R(s, t)$	Volatility of spot rate with time-to-maturity τ at time t given the information at time s
$P(t, \tau)$	Discount bond with time-to-maturity τ at time t
n	Number of cashflows of coupon paying bond
τ	Vector of payment dates of coupon paying bond, $\boldsymbol{\tau} \in \mathbb{R}^n$

c	Vector of coupon payments of coupon paying bond, $\mathbf{c} \in \mathbb{R}^n$
$P(t, \tau, \mathrm{c})$	Coupon paying bond with vector of coupon payments c and coupon payment dates τ at time t
$B(t)$	Money market account at time t
$C(t, K, \tau_1, \tau_2)$	Call option with exercise price K, time-to-maturity τ_1 on discount bond with maturity τ_2 at time t
$P(t, K, \tau_1, \tau_2)$	Put option with exercise price K, time-to-maturity τ_1 on discount bond with maturity τ_2 at time t
N	Number of traded assets
M	Number of states in discrete-time economy
$N(t)$	Vector of portfolio weights or trading strategy at time t
$V_t(N(t))$	Value of trading strategy at time t
I	I Identity matrix, $\boldsymbol{I} \in \mathbb{R}^{M \times M}$
$\boldsymbol{\iota}$	$(1,\ldots,1)^T$, $\boldsymbol{\iota} \in \mathbb{R}^M$

PREFACE

The main reason for starting my Ph.D. study in October 1988 was the expectation of getting the opportunity to investigate thoroughly an economically relevant valuation problem as well as the opportunity to combine practical and theoretical experiences within the financial investment community. After four years of research, consultancy and writing a thesis, these expectations are to a great extent fulfilled.

Working together with Ton Vorst, Angelien Kemna and Peter Schotman was a very learning experience and I am proud of the final result, which could not have been established without their supervision and co-operation.

The two periods of three months I worked for the Shell Pension Fund in The Hague increased my interest in the modelling of financial securities and I am very much indebted for their efforts to initiate the opportunity to work for J.P.Morgan Securities Inc. in New York. In addition, I would like to thank these companies and ABN AMRO for providing valuable data on yield curves, bond prices and interest rates.

Although the writing of a Ph.D. thesis is primarily an individual experience, this work could not have been completed without the help and assistance of my family, friends, and colleagues.

My colleagues at the Tinbergen Institute, and especially those who shared a room with me during the last four years, were of great assistance in establishing this work. It was a pleasure to work at the Tinbergen Institute and the importance of shared research experiences cannot be overestimated. My research also greatly benefited from the visits to scientific conferences and colleagues abroad and I would like to thank the Tinbergen Institute, the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO) and the Erasmus Center for Financial Research for giving me these opportunities.

I am very much indebted to my friends, and especially Willem Brussaard and Peter Serhalawan, for patiently listening to a lot of problems I encountered during the last years. In addition, the a.i.o. soccer team ‘Bahrakkers/Faionoord’ and the social meetings on Thursday night were an important part of my Ph.D. study.

Finally, this preface is very well suited to thank my parents and my brother for accepting my apologies for not visiting them as much as we wanted. I hope this finished thesis serves as a final excuse.

This thesis would definitely not have been completed without the encouragement and patience of Barbara. I thank her for all those evenings I had to work and for her support

and assistance when I worked in New York. I hope to be able to give her back all the love she demonstrated.

Jeroen F.J.de Munnik

1
INTRODUCTION

> If Aristotle is to be believed, Thales started it all in the 6th century BC with an option on olives. But it was only after academic options-pricing theory of the 1960s and 1970s met the volatile financial markets of the deregulating 1980s that options took off. So did futures, warrants, swaps, swaptions, collars, caps, floors, circuses and scores of other products known collectively as derivatives. Nothing today is transforming financial markets as rapidly and completely as what their inventors like to call tools for the management of financial risk. Nothing now gives so many financial regulators so many nightmares.[…] What especially worries most of them is that banks are the biggest traders and counterparts, and neither banks (nor anyone else) understand the risks well enough to price them properly. A derivatives disaster could overwhelm the world's financial system, as third-world debt, highly leveraged transactions and property lending have not managed to do.[1]

This quotation is but one of many examples of the caution and sceptical attitudes existing regarding the development and further evolution of the use of financial derivative securities. At the same time, it stresses the importance and need of an accurate assessment of the values of the different contingent claims and the risks involved.

During the last two decades a lot of academic research has concentrated on the theoretical valuation and the associated empirical validity of commonly known contingent claims like call and put options on stocks. The relatively high risk of stocks in comparison to other alternative financial assets like bonds and the corresponding popularity of these instruments for portfolio management easily explains the primary focus of contingent claim research during this period.[2]

In the last decade, however, increased attention has been paid to the valuation of contingent claims whose values depend on the term structure of interest rates and its subsequent movement over time. Although the level of price risk of Traded Government Bonds[3] may give the impression at first sight of a relatively unimportant problem, the variety of financial instruments with complex option characteristics such as callable bonds, different types of mortgages and the delivery option embedded in a futures contract and the size of the different markets in which these instruments are traded,[4] definitely leads to an opposite conclusion. In addition, the modelling and estimation of the stochastic dynamics of the yield curve not only enables an assessment of the interest rate risk of the above-mentioned instruments but also allows for a general interest rate risk management of fixed income portfolios.

Regarding the strong attention that has been paid to the theoretical valuation problem of ordinary options on stocks, it is important to explain the institutional differences between a stock and a bond to understand and justify the separate treatment of the valuation of interest rate derivative securities.

The main difference between a stock and an ordinary coupon paying bond is the certainty at some valuation date of the amounts and corresponding payment dates of the different coupons and face value. Obviously, this affects the possible price movements of bonds in comparison to those of stocks. Near the final maturity date of a bond, for example, the probability of an increase in value of a par bond is much smaller than it is at some other valuation date, all else being equal. In the case of stocks, however, there is no reason why such particular stochastic behavior can be assumed or derived from institutional characteristics. As another result of this price effect, the corresponding volatility of possible price movements decreases as the maturity of the bond decreases. The range of possible bond prices that can be attained with some probability narrows when the final payment date is reached.

Although not generally empirically justified, one of the basic assumptions in the classical stock options valuation problem is a constant interest rate at which long and short asset positions can be financed. In the case of interest rate derivative securities, however, it is clearly theoretically inconsistent to adopt this assumption. The relationship between bond values and the term structure of interest rates implied by the familiar discounting of future payments, necessitates a formulation of the stochastic movement of the yield curve over time. As will be seen in the remainder of this thesis, this difference results in an increased theoretical and empirical complexity of the valuation problem of these contingent claims and leads to an important distinction between the different valuation methods.

RESEARCH OBJECTIVES

Having explained and justified a separate and extensive treatment of the general valuation problem of interest rate derivative securities, it is important to formulate and discuss some more specific research objectives.

Given the above-mentioned institutional and theoretical differences between the stochastic dynamics of stocks and bonds affecting the valuation problem of interest rate derivative securities, it is both necessary and important to investigate and to give an overview of the different conditions under which derivative securities can be valued. Given the characteristics of some contingent claims, for example, and given the stochastic properties of the underlying values of these claims, is it possible to formulate general conditions regarding the possible values of the claims that exclude riskless arbitrage opportunities between the underlying values and the derivative securities?

The increased academic interest that has been paid during the last few years to the valuation of interest rate derivative securities has resulted in a variety of different theoretical models. To be able to investigate the possible advantages and drawbacks of these approaches with respect to each other and to decide under which theoretical and empirical circumstances a particular model should be preferred, it is necessary to develop a basic classification scheme. Is it possible, therefore, to classify the different interest rate models according to some basic or general characteristics?

The first paragraph of this chapter mentioned that during the last few years a lot of highly complex derivative securities have been developed and introduced. Because of the possible complicated payouts of some of these securities, the actual valuation given some interest rate model relies heavily on the use of numerical methods. Is it possible to classify these numerical approaches also and to develop some decision rules to be able to decide under which conditions a particular method should be preferred?

The principal reason for the increased theoretical attention to the valuation of interest rate derivative securities has been the aim to incorporate the institutional characteristics and the observed empirical properties of interest rate dynamics as much as possible into the derivative securities models. Less emphasis has been paid, however, to an empirical evaluation of the different models and to an assessment of the actual need to incorporate these characteristics. As the resulting valuation complexity and estimation difficulties generally rapidly increase as more properties are built into the model, it is obvious that this comparison should be carried out. Is it possible to distinguish different interest rate models with respect to their empirical validity?

OVERVIEW

The recent extensive theoretical developments within the field of the valuation of interest rate derivative securities implies the necessity of an accurate description of the various models and corresponding numerical approaches according to some basic characteristics. In addition, an empirical investigation has to be carried out to compare different interest rate models and to assess the trade-off between the desire to incorporate as many institutional characteristics as possible and the resulting theoretical and numerical complexity. The different research objectives of this thesis discussed in the previous section are, therefore, either theoretical or empirical in nature and justify a separation of the theoretical and empirical issues into two parts. The distinction further clarifies the clear emphasis on theoretical research during the past and the only recent empirical developments.

The first chapter of the theoretical part (Chapter 2) derives and formulates conditions under which riskless arbitrage opportunities between contingent claims and the corresponding underlying assets are excluded and under which it is possible to determine a unique arbitrage-free value of a derivative security. It is shown that every contingent claim can uniquely be valued in this way if there exists a unique equivalent probability measure such that the stochastic process of the underlying values of this security in terms of a short-term money market account is a martingale. Furthermore, the unique arbitrage-free value of the claim is shown to be equal to the discounted expected value of the payout of this claim under the equivalent martingale measure.

To proceed with the theoretical investigation of the valuation of interest rate derivative securities, Chapter 3 presents a general overview and classification of the different theoretical approaches. The major distinction between the different approaches can be made with respect to the modelling of the underlying values of the different securities.

Just as in the case of the classical stock options valuation problem, the underlying values are explicitly modelled in the direct approach. Given the stochastic dynamics of these values, contingent claims can be valued according to the unique equivalent martingale measure approach. The indirect approach, however, starts with a description of the stochastic dynamics of interest rates. The exclusion of arbitrage opportunities between different bonds, then, results in a description of the specific shape of the term structure of interest rates at a given date and the corresponding stochastic movement over time. The equivalent martingale approach and the derived yield curve then enable the arbitrage-free valuation of any interest rate contingent claim.

Chapter 4 discusses the approach based on the explicit modelling of bond prices. Given the above-mentioned academic research regarding the valuation of options on stocks, it is natural to extend these approaches to incorporate the specific characteristics of bonds and to value contingent claims analogously. Because the model presented basically combines and extends two existing models with respect to the theoretical validity of the proposed stochastic processes, a lot of attention is paid in this chapter to regularity of the processes and the existence of a unique equivalent martingale measure.

The different models within the indirect approach are discussed and illustrated in Chapter 5. Within this class, the different models can be further distinguished according to the term structure of interest rates at the valuation date. The first part of this chapter discusses those models in which the yield curve is endogenously implied by the stochastic characteristics of the interest rate processes and the no-arbitrage relationships. The second part then presents the models in which this term structure is exogenously specified at the given valuation date. In this part, a similar distinction is made between the endogenous and exogenous term structure of interest rate volatilities at some date.

As mentioned above, the complexity of the different valuation models and of the characteristics of the different contingent claims results in the use of numerical valuation methods. In Chapter 6, three approaches are discussed to value a contingent claim numerically given a general stochastic process of the underlying state variable. In addition to this overview, general decision rules are developed to assess and distinguish the numerical accuracy of the different methods in terms of the numerical complexity. Although these methods allow for a numerical approximation of rather general interest rate processes, some interest rate models enable a significant simplification of the original stochastic process with respect to the numerical applicability or can only be approximated by different approaches. At the end of this chapter, some specific numerical methods to value interest rate derivative securities in the case of these interest rate models are presented and further developed.

In the first two chapters of the empirical part of this thesis, the estimation and corresponding results of two models within the class of the endogenous term structure of interest rate models are discussed. The reasons for concentrating on these two models within this particular class are two-fold. The endogenous yield curve at some valuation date is a result of the stochastic characteristics of the underlying interest rate process and the no-arbitrage conditions. A cross-sectional estimation of this yield curve and a time series estimation of the corresponding interest rate process, therefore, enables an interesting comparison between the implicitly and explicitly estimated interest rate processes. In addition, because the two models basically differ with respect to the assumed interest rate process, an empirical comparison allows for an assessment of the increased complexity resulting from the exclusion of negative nominal interest rates.

Chapter 7, therefore, discusses the time series estimation of the two models and the corresponding results. Chapter 8 follows with the estimation method and results of the cross-sectional approach. In addition to an actual comparision of the estimated interest rate processes, these two chapters compare the implications of the estimations for the valuations of European call options on discount bonds.

In the last chapter of this part, some serious problems resulting from the use of principal component analysis to value interest rate derivative securities are discussed. Because this estimation technique is used both by practitioners and academics, Chapter 10, finally, provides an important and illustrative example of taking a cautious approach to the practical implementation of valuation models.

NOTES

1 "Taming the Derivative Beast", 23 May 1992, *The Economist,* pp. 85–6.

2 See, for example, Malkiel (1990, p. 221).

3 These bonds are assumed to be default-free and as such, the risk of an actual default may be ignored. Throughout this thesis, only this particular type of bonds is considered.

4 For an example in the case of the Dutch Fixed Income Market, see de Munnik and Vorst (1988).

Part I
THE THEORETICAL VALUATION OF INTEREST RATE DERIVATIVE SECURITIES

2
ARBITRAGE OPPORTUNITIES AND THE VALUATION OF CONTINGENT CLAIMS

The notion of the existence of arbitrage opportunities in financial markets constitutes a significant area of continuing research in financial economics. Given the current values at which traded securities, such as stocks and bonds, can be bought and sold, is it possible, for example, to construct so-called arbitrage portfolios containing some or all of these securities with a current total value of zero and a strict positive value, with certainty, some time in the future? If it is, one would not expect to achieve an equilibrium between or within financial markets. Investors would recognize these opportunities and perform arbitrage strategies on a very large scale by buying securities that are relatively undervalued and vice versa. This will change the prices of traded securities and eliminate the opportunities for arbitrage, accordingly. An interesting problem arising from this mechanism is the formulation of conditions regarding the current values of traded securities and their probability distributions so that such arbitrage opportunities do not exist.

Another interesting problem, related to the above-mentioned questions, is the valuation of contingent claims or derivative securities, securities whose values depend on the prices and characteristics of one or more traded securities. Is it possible to consider these claims as portfolios of traded securities and to replicate the payout of these claims by trading in the underlying securities? Can unique prices or a range of prices for these claims be derived such that arbitrage opportunities between the claims and other securities are eliminated?

This chapter deals with the formulation and derivation of general conditions under which the unique value of a contingent claim can be derived. It will show the existence of an arbitrage-free price of a contingent claim to be equivalent to the existence of a unique equivalent probability measure, under which prices of traded securities relative to a short-term money market account are martingales. Although this thesis is mainly concerned with the valuation of interest rate derivative securities, this chapter focuses on the valuation of derivative securities in a more general way. As will be explained in later chapters, the results derived are shown to be readily applicable to the valuation of specific interest rate contingent claims.

The first part of this chapter is concerned with economies in which prices of securities change randomly at discrete points in time and the sample space is finite. This relatively simple structure allows for explanation of the concepts and definitions of arbitrage opportunities, trading strategies and contingent claim valuation in simple mathematical terms that facilitate economic interpretation. As an illustration of the main theorems of this section, an example is provided: how the value of a European call option on a discount bond is derived. The second part of this chapter will then extend this simple economy to an economy in which prices change continuously and investors are allowed to trade continuously. It will be shown that the results and main conclusions of the first part of this chapter are directly extendable to these continuous-time economies. Because the valuation technique of contingent claims derived in this section may seem rather technical and abstract, its practical relevance will be demonstrated with an example in which the familiar Black and Scholes (1973) formula for the value of a European call option on a non-dividend paying stock is derived in detail.

THE DISCRETE TIME CASE

This section focuses on economies in which prices change randomly at discrete points in time, the sample space is finite, and investors are allowed to trade in these securities at these discrete points in time. For expository reasons, the section starts with a one-period economy. By concentrating first on this simple economy, problems related to the rebalancing of portfolios and the comparison of securities with differing maturity times are avoided. After this, the derived results are extended to a multi-period economy.

One-Period Economies

Consider an economy with only two dates, a current trading date 0 and a final date T, $T \geq 0$. The Probability space $(\Omega, \mathcal{F}, P)$ is specified and fixed as follows. The sample space Ω has a finite number of elements ωi, j=1,…, M, each of which can be interpreted as a possible state of the world. All probability measures P are equivalent in requiring that the probability measure P may be replaced by another equivalent measure P^*. $P(\omega j) > 0$ for all ωj Ω,j=1,…, M, meaning that investors have to agree on which states of the world are possible. It is not necessary that investors agree on the actual assessment of probabilities of states of the world at time T. Revelation of information through time is specified by a filtration $F = \{\mathcal{F}_0, \mathcal{F}_T\}$, where it is assumed, without loss of generality, that $\mathcal{F}_0$ equals the trivial tribe {Ø, Ω} and that $\mathcal{F}_T$ equals $\mathcal{F}$, the set of all subsets of Ω

There are N traded securities or marketed claims at time 0, of which the prices are given by the N-dimensional vector $S(0)$, with component prices or values $S1(0), S2(0),\ldots, S_N(0)$. Each component process S_i is strictly positive and adapted to the filtration F, reflecting the assumption of limited liability and implying that investors know at time 0 and time T the values of all traded securities. The set of possible values for these securities at time T is specified by a matrix $S(T) \in \mathbb{R}^{N \times M}$, so $S(T)_{i,j} = S_i(T)(\omega_j)$ for i=1,…, N and j= 1,…, M. Without loss of generality, the first asset is assumed to be a riskless bond with a current value of 1 and paying an interest rate r(0), r(0) ≥ 0, or

$$S(T)_{1,j} = (1 + r(0)) \qquad \forall j = 1, \ldots, M \tag{2.1}$$

This assumption implies that investors are able to invest in a riskless money market account, which will facilitate the discussion of well-known interest rate models to be treated in later chapters.

A portfolio of traded assets or a trading strategy is defined as a predictable vector process *N(T)* with components $N_1(T), N_2(T), \ldots N_N(T)$, implying $N(T) \in \mathcal{F}_0$; $\mathcal{N}$ is denoted as the set of trading strategies. Each component *Ni(T)* can be interpreted as the number or quantity of security i,i= 1,…, *N* in this portfolio, which may be positive or negative. The requirement that the vector process *N(T)* must be predictable, is merely a stipulation that the portfolio has to be established before the announcement of the prices of the traded securities at time *T*. The value $V_0(N(T))$ of the portfolio at time 0 is now

$$V_0(N(T)) = S(0)^T N(T) \tag{2.2}$$

whereas the possible values at time *T* can be expressed as the vector

$$V_T(N(T)) = S(T)^T N(T) \tag{2.3}$$

The introduction to this chapter raised the question of the existence of arbitrage opportunities. To be able to answer this question, one must be able to specify the conditions under which there are riskless arbitrage opportunities. In this study, as in Ingersoll (1987, p. 53), an arbitrage opportunity of the first type exists if there is a portfolio *N(T)* such that the current value of the portfolio or initial investment is zero and the value of the portfolio at the final date *T* is non-negative with probability one and strictly positive with positive probability. Formally stated,

$$V_0(N(T)) = 0 \tag{2.4}$$

and

$$V_T(N(T)) \geq 0 \qquad V_T(N(T)) \neq 0 \tag{2.5}$$

In relation to this concept of arbitrage and as a useful concept for multiperiod economies, an arbitrage opportunity of the second type is defined to exist if there is a portfolio *N(T)* such that the current value of the portfolio is negative with probability one and the final value is non-negative, that is,

$$V_0(N(T)) < 0 \tag{2.6}$$

and

$$V_T(N(T)) \geq 0 \tag{2.7}$$

The following lemma states that if an arbitrage opportunity of the second type exists, an arbitrage opportunity of the first type exists, too.

Lemma 2.1 *If there exists an arbitrage opportunity of the second type in the one-period economy, then there also exists an arbitrage opportunity of the first type.*

Proof Suppose $N^2(T)$ is an arbitrage opportunity of the second type, having an initial negative investment of $V0(N^2(T))$. Create a portfolio $N^1(T)$ equal to $N^2(T)$ and an amount

$-V0(N^2(T))$ invested in the riskless money market account, making the initial investment of this portfolio zero. So,

$$N_1^1(T) := N_1^2(T) - V_0(N^2(T))$$

and

$$N_i^1(T) := N_i^2(T) \qquad \forall i = 2, \ldots, N$$

The possible values of the portfolio at the final trading date *T* are

$$V_T(N^1(T)) = V_T(N^2(T)) -$$

$$(V_0(N^2(T))(1 + r(0)), \ldots, V_0(N^2(T))(1 + r(0))^T >$$

$$V_T(N^2(T)) \geq 0$$

As a result, portfolio $N^1(T)$ creates an arbitrage opportunity of the first type, which completes the proof.■

An economy containing arbitrage opportunities cannot be in equilibrium. Investors would recognize these opportunities immediately and construct such portfolios. Securities that are relatively overvalued are sold and vice versa. This will change the prices of these securities and arbitrage opportunities will cease to exist. Given the prices of the traded securities at time 0, S(0), and given the matrix of possible prices at the final trading date *T*, *S(T)*, is there a simple condition that excludes the possibility of riskless arbitrage profits? To show that this is indeed the case, one starts with the following lemma, stating the equivalence between the existence of arbitrage opportunities of the first type in this economy and in a corresponding economy where prices are expressed in terms of the first security or value of the money market account, that is,

$$\begin{aligned} S^*(0) &= S(0) & S^*(0) &\in \mathbb{R}^N \\ S^*(T) &= (1 + r(0))^{-1} S(T) & S^*(T) &\in \mathbb{R}^{N \times M} \end{aligned} \qquad (2.8)$$

As a direct result, the exclusion of arbitrage opportunities of the first type in this relative economy implies the exclusion of the same opportunities in the original economy. Although no economic justification for this transformation is available at this moment, it will be seen that the interpretation of subsequent results is greatly facilitated.

Lemma 2.2 *There exists an arbitrage opportunity of the first type in the one-period economy if and only if there exists an arbitrage opportunity of the first type in the corresponding one-period economy, in which prices are expressed in terms of the value of the riskless money market account.*

Proof Suppose an arbitrage opportunity of the first type exists in the original one-period economy, that is, there exists an *N(T)* such that

$$S(0)^T N(T) = 0$$

and

$$S(T)^T N(T) \geq 0 \qquad S(T)^T N(T) \neq 0$$

Executing the same trading strategy *N(T)* in the discounted economy, gives

$$S^*(0)^T N(T) = 0$$

and

$$S^*(T)^T N(T) \geq 0 \qquad S^*(T)^T N(T) \neq 0$$

which creates an arbitrage opportunity of the first type. Because a proof of the converse statement is completely analogous, it is omitted.■

The following theorem shows that if and only if there exists an equivalent probability measure such that relative prices are martingales, or equivalently, such that relative values of trading strategies are martingales, riskless arbitrage opportunities of the first type in this relative economy do not exist. Combining this result with the lemmas above, one can conclude that arbitrage opportunities in our one-period economy do not exist, either.

Theorem 2.3 *Arbitrage opportunities of the first type in the relative oneperiod economy do not exist if and only if there exists a probability measure equivalent to P, such that prices are martingales with respect to this measure.*

Proof Suppose arbitrage opportunities of the first type do not exist. First construct the following linear programming problem:

$$\begin{aligned} &Min\, S^*(0)^T N(T) + 0^T x \\ &S.t.\, S^*(T)^T N(T) - Ix \;\; \geq 0 \\ &\qquad\qquad\qquad\qquad \iota^T x = 1 \\ &\qquad\qquad\qquad\qquad\; x \geq 0 \end{aligned}$$

with $I \in \mathbb{R}^{M \times M}$, and $\iota \in \mathbb{R}^M$. This problem is an extension of the problem formulated in Ingersoll (1987, p. 55), in which the equivalence between the exclusion of arbitrage opportunities of the second type in the relative economy and the existence of a probability measure such that relative prices are martingales is proved. However, as is clear from Lemma 2.1 above, the exclusion of arbitrage opportunities of the second type does not necessarily imply the exclusion of arbitrage opportunities of the first type. To prohibit these opportunities also, the probability measure has to be equivalent to *P*.

The objective function of this linear programming problem can be interpreted as the initial value of the portfolio. The first constraint denotes the value of this portfolio in all possible states of the world at time *T*. Combined with the other two constraints, this value is non-negative in all possible states and strictly positive in at least one state. By taking $N(T) = (1/S_1^*(0), 0, \ldots, 0)^T$ and $x = \iota/M$, one sees that the problem is feasible and has a minimum value of the objective function that is strictly positive.[1] Now comes the formulation of the corresponding dual linear programming problem:

$$\begin{aligned} Max\, y + \quad & 0^T q \\ S.t.\ \iota y - \quad & Iq \leq 0 \\ & S^*(T)q = S^*(0) \\ & q \geq 0 \end{aligned}$$

Because of the feasibility of the primal linear programming problem, this corresponding dual problem is known to be feasible and to have a maximum value of the objective function that is strictly positive. This implies that the value of *y*, for which this optimum is obtained, is strictly positive, too, causing the vector *q* to be strictly positive. Because of the first equality in the formulation of the problem above, equation (2.1) and the definition of *S*(T)*, this is also true $\iota^T q = 1$

The vector is thus suitable as an equivalent probability measure. The remaining equalities above show that under this measure, all traded security prices are martingales.

To prove the converse statement, we simply reverse the different steps above. The primal linear programming problem has a minimum value of the objective function that is strictly positive, excluding arbitrage opportunities of the first type.[2]■

$\mathcal{Q}$ is now defined as the set of equivalent martingale measures. The elements $Q \in \mathcal{Q}$, each of which can be represented by an *M*-dimensional vector *q*, with $q > 0$ and $i^T q = 1$, are often called 'risk-neutral probabilities', because the expected return on all traded securities with respect to these measures is equal to the riskless interest rate, r(0). This does not mean, however, that investors have to be risk-neutral to exclude arbitrage opportunities. Theorem 2.3 implies only that there has to exist a probability measure under which investors are risk-neutral.

We define a contingent claim as a random variable X_T on the probability space $(\Omega, \mathcal{F}, P)$. This random variable can be interpreted as a contract or agreement paying $X_T(\omega_j)$ at time *T* if state ω_j Ω, $j = 1,\ldots, M$ pertains. Let $\mathcal{X}$ denote the set of all such claims. A contingent is here attainable if there exists a trading strategy that generates X_T, that is, there is a $N(T) \in \mathcal{N}$ such that $VT(N(T)) = X_T$. The initial value $V_0(N(T))$ of such a replicating portfolio can then be regarded as the price π of this attainable claim. Is this price unique or, in other words, are there different trading strategies generating X_T with different initial values? The following theorem states that in an economy in which arbitrage opportunities do not exist, prices of attainable contingent claims are indeed unique.

Theorem 2.4 *If arbitrage opportunities do not exist in the one-period economy, there is a single price* π *associated with each attainable contingent claim* X_T*, which satisfies* $\pi = E_Q\left(\frac{1}{1+r(0)} X_T | \mathcal{F}_0\right)$.

Proof Define $\tilde{\mathcal{N}} \subseteq \mathcal{N}$ as the set of trading strategies generating the contingent claim X_T. Because $\mathcal{Q}$ is not empty,

$$E_Q\left(\frac{1}{1+r(0)} X_T | \mathcal{F}_0\right) = E_Q\left(\frac{1}{1+r(0)} V_T(N(T)) | \mathcal{F}_0\right) =$$

$$q^T \frac{S(T)^T}{1+r(0)} N(T) = S(0)^T N(T) = V_0(N(T)) = \pi$$

for all $Q \in \mathcal{Q}$ and $N(T) \in \tilde{\mathcal{N}}$. To show that π is the unique price associated with the contingent claim X_T. suppose there are $N_1(T), N_2(T) \in \tilde{\mathcal{N}}$, with corresponding price π_1 and π_2, respectively, generating X_T. Taking the difference results in

$$0 = E_Q\left(\frac{1}{1+r(0)} V_T(N_1(T)) - \frac{1}{1+r(0)} V_T(N_2(T)) | \mathcal{F}_0\right)$$

$$= E_Q\left(\frac{1}{1+r(0)} V_T(N_1(T)) | \mathcal{F}_0\right) - E_Q\left(\frac{1}{1+r(0)} V_T(N_2(T)) | \mathcal{F}_0\right)$$

$$= V_0(N_1(T)) - V_0(N_2(T)) = \pi_1 - \pi_2$$

This completes the proof.■

From this theorem it is known that if a contingent claim $X_T \in \mathcal{X}$ is attainable, a unique price for this claim can be computed or derived. But how does one know if a contingent claim is attainable? Under which circumstances concerning this one-period economy is every contingent claim attainable? If, in fact, every contingent claim is attainable, our economy or security market model can be judged to be complete. To avoid any trivial complications, first a non-degeneracy condition is imposed. This economy or price process contains a redundancy if there exists a trading strategy $N(T) \in \mathcal{N}, N(T) \neq 0$ such that $V_T(N(T))=0$. If such redundancy existed, possession of some security would be completely equivalent to a portfolio of other traded securities. Ignoring this redundant security would therefore not limit the ability to attain contingent claims, which justifies this assumption. The following theorem shows the condition under which this market model contains no arbitrage opportunities and is complete.

Theorem 2.5 *The one-period economy contains no arbitrage opportunities and is complete if and only if there exists a unique equivalent probability measure such that relative prices are martingales.*

Proof Because no redundancy exists, *Rank(S(T))=N.* For this model also to be complete, it must be stipulated that for every contingent claim $X_T \in \mathcal{X}$, there exists a trading strategy $N(T) \in \mathcal{N}$ such that $X= V_T(N(T))=S(T)^T N(T)$, implying *Rank(S(T))=M* and therefore *S(T)* to be invertible. The exclusion of arbitrage opportunities is equivalent to the set of equivalent martingale measures $\mathcal{Q}$ being not empty. The invertibility of *S(T),* then, implies that $\mathcal{Q}$ is a singleton.

To prove the converse, because it is known that the existence of an equivalent martingale measure implies no arbitrage opportunities and because this measure is unique, we have *Rank(S(T))=Rank(S*(T))=M.* The no-redundancy assumption then completes the proof.■

Based on this theorem, it is known that every possible contingent claim is attainable, once the existence of a unique equivalent martingale measure is determined. According to Theorem 2.4 then, a claim's unique price is equal to its discounted expected value at maturity, where the expectation has to be taken with respect to the equivalent martingale measure.

Multi-Period Economies

In this section, the one-period economy is extended to a multi-period economy. In this economy, security prices change randomly at discrete points in time, the sample space is again finite and investors can trade in the securities at these discrete points in time. Although most of the concepts and theorems derived above are easily applicable, some interesting problems are worth mentioning. Contrary to the economy described above, for example, contingent claims can reach maturity at different or even random times. Related to this, securities and contingent claims can pay dividends or have other payouts at certain points in time and a refinement and modification of trading strategies must be made to be able to generate those contingent claims. As will be shown, however, the unique arbitrage-free valuation of every contingent claim is again equivalent to the existence of a unique equivalent martingale measure, a result which will be illustrated by the valuation of a European call option on a discount bond.

The characterization of the multi-period economy starts with a finite set of trading dates $\mathcal{T} = \{0, 1, \dots, T\}$, with 0 the current trading date and *T* the final date. The probability space $(\Omega, \mathcal{F}, P)$ is specified and fixed as in our one-period economy. The sample space Ω has a finite number of elements and the probability measure *P* may be replaced by an equivalent probability measure *P** without changing the conclusions. The filtration $F = \{\mathcal{F}_0, \dots, \mathcal{F}_T\}$ is specified as an increasing set of σ-algebras, that is, $\mathcal{F}_s \subseteq \mathcal{F}_t$ for all $s, t \in \mathcal{T}, s \leq t$. $\mathcal{F}_0$ equals the trivial tribe {Ø, Ω} and $\mathcal{F}_T$ equals again $\mathcal{F}$, the set of all subsets of Ω.

The prices of the *N* traded securities are specified as an N-dimensional stochastic process $S = \{S(t), t \in \mathcal{T}\}$, with component stochastic processes $S_1(t), \dots, S_N(t)$ that are strictly positive and adapted to the filtration *F*.Because of the requirement that the process *S(t)* be measurable with respect to $\mathcal{F}_t,\ t \in \mathcal{T}$, the investors know at time t the prices of all traded securities at and before time *t*. Again the assumption is made that the first asset is a locally riskless money market account, defined by

$$S_1(0) = 1$$
$$S_1(t) = \prod_{s=0}^{s=t-1} (1 + r(s)) \qquad \forall t \in \mathcal{T} \setminus \{0\} \tag{2.9}$$

The specification of this money market account implies only at this moment that investors have the opportunity at any time t, $t \in \mathcal{T} \setminus \{T\}$, to invest in an account earning a positive riskless rate of *r(t)* during the period [*t*, *t*+1]. As such, the rate of return on this account during a longer period $[t, T'],\ T' \in \mathcal{T},\ T' > t + 1$ is unknown beforehand, as short-term interest rates generally are assumed to be stochastic. In the rest of this section, however, conditions will be derived under which investors are, in fact, able at some time *t* to invest or create trading strategies yielding a riskless return during any period.

A portfolio or trading strategy is defined as a vector process $N = \{N(t), t \in \mathcal{T}\}$ with components $N_1(t),\ldots,N_N(t)$. This process is again assumed to be predictable, meaning $N(t) \in \mathcal{F}_{t-1}$, for all $t \in \mathcal{T} \setminus \{0\}$, because the portfolio *N(t)* has to be established before the announcement of the security prices *S(t)*. The value of the portfolio can now be expressed as a stochastic process $V = \{V_t(N(t)), t \in \mathcal{T} \setminus \{0\}\}$, which is $\mathcal{F}_t$-measurable and has initial value

$$V_0(N(1)) = S(0)^T N(1) \tag{2.10}$$

The value at time *t* is

$$V_t(N(t)) = S(t)^T N(t) \tag{2.11}$$

A general trading strategy may require the addition of new funds after time zero or the withdrawal of funds for consumption. If no funds are withdrawn or added, a portfolio may be called self-financing, implying that changes in the portfolio holdings obey

$$S(t)^T N(t) = S(t)^T N(t+1) \qquad \forall t \in \mathcal{T} \setminus \{0, T\} \tag{2.12}$$

Let $\mathcal{N}$ denote the set of self-financing trading strategies. As in the oneperiod economy, an arbitrage opportunity of the first type exists if there is a self-financing trading strategy *N* such that

$$V_0(N(1)) = 0 \tag{2.13}$$

and

$$V_T(N(T)) \geq 0 \qquad V_T(N(T)) \neq 0 \tag{2.14}$$

An arbitrage opportunity of the second type is defined to exist now if there is a self-financing trading strategy *N* such that

$$V_0(N(1)) < 0 \tag{2.15}$$

and

$$V_T(N(T)) \geq 0 \tag{2.16}$$

These two types of arbitrage opportunities are very general and include all possible arbitrage situations. If, for example, a portfolio exists with zero initial value and a non-negative value at time $t,\ t \in \mathcal{T} \setminus \{T\}$, this value will be invested at time *t* in the money market account. The value of the portfolio at the final trading date *T*, then, is non-negative too, creating an arbitrage opportunity of the first type. A similar argument can be used to show that a strategy with an initial zero investment and final non-negative value, yielding positive payments at some trading dates, is also an arbitrage opportunity of the first type. The following lemma states a familiar relationship between the two types of arbitrage.

Lemma 2.6 *If there exists an arbitrage opportunity of the second type in the multi-period economy, then there also exists an arbitrage opportunity of the first type.*

Proof Assume, as in the proof of lemma 2.1, that there is a trading strategy that is an arbitrage opportunity of the second type. Because the initial value of this portfolio is negative, this cash inflow is invested in the money market account, creating an alteraative portfolio with an initial value of zero. Because the arbitrage portfolio has a final value that is non-negative, the alternative portfolio will have a final value that is non-negative and not equal to zero, because of the initial investment in the money market account. This alternative portfolio is then an arbitrage opportunity of the first type.■

To facilitate the interpretation of the theorems concerning the exclusion of arbitrage opportunities and completeness of the security market as in the one-period economy, prices of traded securities are expressed in terms of the value of the money market account, and the resulting $\mathcal{F}_t$-measurable price process $S^* = \{S^*(t),\ t \in \mathcal{T}\}$ is defined as

$$\begin{aligned} S_i^*(0) &:= S_i(0) \qquad \forall i = 1, \ldots, N \\ S_i^*(t) &:= \prod_{s=0}^{s=t-1} (1 + r(s))^{-1} S_i(t) \quad \forall i = 1, \ldots, N \quad \forall t \in \mathcal{T} \setminus \{0\} \end{aligned} \tag{2.17}$$

The stochastic process of the value of trading strategies $V^* = \{V_t^*(N(t)),$ $t \in \mathcal{T} \setminus \{0\}\}$ is defined accordingly and this multi-period economy will be referred to as the relative multi-period economy.

Lemma 2.7 *There exists an arbitrage opportunity of the first type in the multi-period economy if and only if there exists an arbitrage opportunity of the first type in the corresponding multi-period economy, in which prices are expressed in terms of the value of the locally riskless money market account.*

Proof See the proof of lemma 2.2.■

The proof of the following theorem, in which a condition prohibiting arbitrage opportunities in the relative multi-period economy is derived, directly depends on the results obtained in the previous section. By presenting the multi-period economy in the way described above, it will be seen that the multi-period economy is a straightforward extension of the one-period economy.

Theorem 2.8 *Arbitrage opportunities of the first type in the relative multiperiod economy do not exist if and only if there exists a probability measure Q, equivalent to P, such that values of self-financing trading strategies are martingales with respect to this measure.*

Proof Let $\mathcal{P}_t$ denote the partition of Ω underlying $\mathcal{F}_t$, $t \in \mathcal{T}$. Call the discounted prices in the cells in $\mathcal{P}_{t+1}$, which are contained in $A \in \mathcal{P}_t$, together with the discounted prices in *A*, a relative sub one-period economy. This demonstrates the equivalence between the existence of arbitrage opportunities of the first type in the relative multi-period economy and the existence of an arbitrage opportunity in one of the relative sub one-period economies. Combining this equivalence with the results of Theorem 2.3, the proof is complete.

Suppose an arbitrage opportunity $N \in \mathcal{N}$ of the first type exists in the relative multi-period economy and not in any of the relative sub oneperiod economies. It will be shown by induction that the value of this portfolio is zero in all possible states of the world at time *T*, which is a contradiction.

At time 0, the value of the arbitrage portfolio is, because $N \in \mathcal{N}$,

$$V_0^*(N(1)) = 0$$

By assumption the first relative sub one-period economy does not contain an arbitrage opportunity, so, for all cells $A \in \mathcal{P}_1$,

$$V_1^*(N(1)) = 0$$

The trading strategy is self-financing, which also implies, for all cells $A \in \mathcal{P}_1$,

$$V_1^*(N(2)) = 0$$

Suppose now that at time t, $t \in \mathcal{T} \setminus \{T\}$ there, for all cells $A \in \mathcal{P}_t$

$$V_t^*(N(t+1)) = 0$$

Each of these cells is the initial trading date of a relative sub one-period economy. Therefore

$$V^*_{t+1}(N(t+1)) = 0$$

for all cells $A \in \mathcal{P}_{t+1}$ which are contained in $\mathcal{P}_t$. Because the zero-probabil-ity events do not exist by assumption, the value of the portfolio is zero at time t+1 for all cells $A \in \mathcal{P}_{t+1}$. Because N is self-financing

$$V^*_{t+1}(N(t+2)) = 0$$

for all cells $A \in \mathcal{P}_{t+1}$ and $t+2 < T$.

To prove the converse statement, suppose that there is a relative sub one-period economy with initial trade date $t',\ t' \in \mathcal{T} \setminus \{T\}$ in which an arbitrage opportunity of the first type exists. Suppose, also, that the starting point of this relative sub one-period economy is $A \quad P_t$ and that N' are the portfolio holdings creating this arbitrage opportunity. We then create the following trading strategy,

$$
\begin{array}{lll}
N(t) = 0 & \forall t < t' & \\
N(t) = 0 & \forall t \geq t' & \omega \notin A \in \mathcal{F}_t \\
N(t) = N' & t = t' + 1 & \omega \in A \in \mathcal{F}_{t'} \\
N_1(t) = V^*_{t'}(N(t'+1)) & \forall t > t' & \omega \in A \in \mathcal{F}_t \\
N_i(t) = 0 & \forall t > t' & \omega \in A \in \mathcal{F}_t \quad \forall i = 1, \ldots, M
\end{array}
$$

The portfolio holdings are equal to zero except when $A \in \mathcal{P}_t$ is reached. Then the portfolio holdings are equal to N'. At the end of the period the proceedings are invested in the first asset until the final trading date T. It is easy to verify that this strategy is self-financing and creates an arbitrage opportunity of the first type.■

The following corollary states that under this equivalent probability measure $\mathcal{Q}$, discounted or relative prices of traded securities are also martingales. Although this result is quite obvious in the discrete multi-period economy, and may even seem a little superfluous, it is by no means in the continuous-time economy to be discussed in the next section.

Corollary 2.9 *The relative values of self-financing trading strategies are martingales with respect to the probability measure* $\mathcal{Q}$ *if and only if relative prices of traded securities are martingales with respect to the probability measure* $\mathcal{Q}$.

Proof Suppose relative prices are martingales with respect to the probability measure $\mathcal{Q}$. Using the law of iterated conditional expectations and the predictability of trading strategies, there is for all $N \in \mathcal{N}$ and $s, t \in \mathcal{T}$, $s < t$,

$$E_Q\left(\frac{1}{S_1(t)}V_t(N(t))|\mathcal{F}_s\right) = E_Q\left(\frac{1}{S_1(t)}S(t)^T N(t)|\mathcal{F}_s\right) =$$

$$E_Q\left(E_Q\left(\frac{1}{S_1(t)}S(t)^T N(t)|\mathcal{F}_{t-1}\right)|\mathcal{F}_s\right) = E_Q\left(\frac{1}{S_1(t-1)}S(t-1)^T N(t)|\mathcal{F}_s\right) =$$

$$E_Q\left(\frac{1}{S_1(t-1)}S(t-1)^T N(t-1)|\mathcal{F}_s\right) = \ldots = \frac{1}{S_1(s)}V_s(N(s))$$

To prove the converse, simply construct a self-financing trading strategy by holding a long position in one of the N traded assets.■

Let $\mathcal{Q}$ again denote the set of equivalent probability measures, such that discounted prices are martingales. In the multi-period economy, let $\mathcal{X}$ denote the set of all $\mathcal{F}_t$-measurable contingent claims $\{X_t, t \in \mathcal{T}\}$, where X_t is a contingent claim yielding $X_t(\omega)$ if $\omega \in \mathcal{F}_t$ pertains at time t, and zero otherwise. At first sight, it seems that this set of claims does not contain claims with random maturities or payouts during maturity of the claim. However, because an investor can always invest or withdraw money from the money market account, this set is not too limited for the purpose at hand. Call a contingent claim $X_t, \ t \in \mathcal{T}$ to be attainable if there exists a trading strategy $N \in \mathcal{N}$ which generates Xt, that is, vt,(N(t)) = Xt. Then, call $\pi(s) = V_s(N(s)), \ s \in \mathcal{T}, \ s \leq t$ the value or price of the claim at time *s*. The following theorem is equivalent to Theorem 2.4 and it states the uniqueness of the price of a derivative security given the exclusion of arbitrage opportunities in the multi-period economy.

Theorem 2.10 *If arbitrage opportunities do not exist in the multi-period economy, there is a single price π(s) associated contin with any attainable gent claim* $X_t \in \mathcal{X}, s \leq t, t \in \mathcal{T}$, which satis*fi*es $\pi(s) = E_Q\left(\frac{S_1(s)}{S_1(t)}X_t|\mathcal{F}_s\right)$.

Proof Define $\mathcal{N} \subseteq \mathcal{N}$ as the set of trading strategies generating the contingent claim $X_t \in \mathcal{X}$. Because arbitrage opportunities do not exist, the set is not empty. According to Corollary 2.9 then,

$$E_Q\left(\frac{1}{S_1(t)}X_t|\mathcal{F}_s\right) = E_Q\left(\frac{1}{S_1(t)}V_t(N(t))|\mathcal{F}_s\right) =$$

$$\frac{1}{S_1(s)}V_s(N(s)) = \frac{1}{S_1(s)}\pi(s)$$

for all $N \in \tilde{\mathcal{N}}, s < T$. To show that π(s) is the unique price associated with the contingent claim $X_t \in \mathcal{X}$, suppose there are $N_1, N_2 \in \tilde{\mathcal{N}}, N_1 \neq N_2$, with corresponding prices $\pi_1(s)$ and $\pi_2(S)$, respectively, generating X_t. Taking the difference results in

$$0 = E_Q\left(\frac{1}{S_1(t)} V_t(N_1(t)) - \frac{1}{S_1(t)} V_t(N_2(t)) | \mathcal{F}_s\right) =$$

$$E_Q\left(\frac{1}{S_1(t)} V_t(N_1(t)) | \mathcal{F}_s\right) - E_Q\left(\frac{1}{S_1(t)} V_t(N_2(t)) | \mathcal{F}_s\right) =$$

$$\frac{1}{S_1(s)} V_s(N_1(s)) - \frac{1}{S_1(s)} V_s(N_2(s)) = \frac{1}{S_1(s)} \pi_1(s) - \frac{1}{S_1(s)} \pi_2(s)$$

which completes the proof.■

In this multi-period economy in which investors can trade at different points in time and face a lot of investment opportunities, it is interesting to examine which claims are attainable. The multi-period economy will be complete if every claim in $\mathcal{X}$ is attainable; this is the case if the set of equivalent martingale measures is a singleton.

Theorem 2.11 *The multi-period economy contains no arbitrage opportunities and is complete if and only if there exists a unique equivalent probability measure such that relative prices are martingales.*

Proof The equivalence between completeness of the relative multi-period economy and completeness of all possible relative sub one-period econ-omies, together with Theorems 2.5 and 2.8, is sufficient to prove this theorem. Because a proof of this equivalence is very similar to the proof of Theorem 2.8, it is omitted.■

This last theorem can be used to determine whether a contingent claim is attainable or not. If it is, Theorem 2.10 states that the arbitrage-free price of this claim is unique and equal to its discounted expected value under the equivalent martingale measure. Suppose the multi-period economy is complete and arbitrage opportunities are not possible. How can this valuation procedure be used to determine the value at time 0 of a European call option with maturity $T_1 \in \mathcal{T}$ and exercise price $K < 1$, where the underlying value is a riskless discount bond with face value 1 and maturity $T_2 \in \mathcal{T}, T_1 < T_2$? First, the value $B(T_1, T_2 | \mathcal{F}_{T_1})$ of this discount bond at maturity of the option given the set of information available at this time is derived. Because the bond is default-free, its value at time T_2 is unity and according to Theorem 2.10

$$B(T_1, T_2 | \mathcal{F}_{T_1}) = E_Q\left(\frac{S_1(T_1)}{S_1(T_2)} | \mathcal{F}_{T_1}\right) = E_Q\left(\prod_{s=T_1}^{s=T_2-1} (1 + r(s))^{-1} | \mathcal{F}_{T_1}\right) \quad (2.18)$$

Given the value of the bond at time T_1 and the information available, the option value $C(T_1 | \mathcal{F}_{T_1})$ is simply

$$C(T_1|\mathcal{F}_{T_1}) = Max(B(T_1, T_2|\mathcal{F}_{T_1}) - K, 0) \tag{2.19}$$

The last step in the valuation procedure is again the application of Theorem 2.10 to derive the option's value at time 0. So

$$C(0) = E_Q\left(\frac{S_1(0)}{S_1(T_1)} C(T_1|\mathcal{F}_{T_1})|\mathcal{F}_0\right) \tag{2.20}$$

which can be written after some calculations as

$$C(0) = E_Q\left(\prod_{s=0}^{s=T_2-1}(1+r(s))^{-1}|\mathcal{F}_0 \cup \{B(T_1, T_2) > K\}\right) - \tag{2.21}$$

$$KB(0, T_1|\mathcal{F}_0)Pr_Q(B(T_1, T_2) > K|\mathcal{F}_0)$$

where $B(0, T_1|\mathcal{F}_0)$ denotes the value at time 0 of a riskless bond maturing at time T_1, and $Pr(B(T_1, T_2) > K|\mathcal{F}_0)$ the probability at time 0 of $\{B(T_1, T_2) > K\}$ under the equivalent probability measure Q.

CONTINUOUS-TIME ECONOMIES

In the previous section, the one-period economy was extended to the multi-period economy, in which prices change randomly at discrete points in time and investors can trade in the securities at these discrete points in time. This section will discuss the valuation of derivative securities and the exclusion of arbitrage opportunities in a continuous-time economy. This economy, in which prices change randomly and continuously, and in which investors can modify their portfolios of traded assets continuously, is the framework of a lot of financial research, from which the derivation of the value of a European call option on a stock by Black and Scholes (1973) is a well-known example.

This section starts with a general description of the stochastic movement of prices of traded securities that is similar to Harrison and Pliska (1981). The equivalence between arbitrage opportunities in this economy and an economy where prices are taken relative to a locally riskless money market account will be demonstrated, and it will be shown, as in the multi-period economy, that the security market contains no arbitrage opportunities and that every contingent claim is attainable if there exists an equivalent probability measure such that the relative value of a self-financing portfolio is a martingale. As such, this result is similar to the one obtained in Theorem 2.11 of the previous section. However, as will become clear later on, the equivalence between the martingale property of the relative value of self-financing portfolios and the relative value of traded securities under an equivalent probability measure no longer holds, resulting in additional restrictions on the portfolio weights. We illustrate these results by an example in which we derive the Black-Scholes option pricing formula.

In the continuous-time economy, there is an initial trading date 0 and a fixed planning horizon *T*. Given this planning interval, the continuoustime uncertainty is specified by the following filtered probability space $(\Omega, \mathcal{F}, F, P)$. In this probability space, Ω denotes the state space, *P* a

probability measure and *F* the filtration of increasing sub-*σ*-algebras $\mathcal{F}_t$, $0 \leq t \leq T$, which satisfy the following usual conditions:[3]

$$\mathcal{F}_0 = \{A \subset \Omega | P(A) = 0\} \cup \Omega$$

$$\mathcal{F}_t = \cap_{s>t} \mathcal{F}_s$$

$$\mathcal{F}_T = \mathcal{F}$$

As in the discrete case, investors have only to agree on the null sets of the probability measure instead of on an actual assessment of probabilities of certain events, which implies that *P* can be replaced by any equivalent probability measure *P**.

Define the stochastic process $S=\{S(t), 0 \leq t \leq T\}$ to be a real-valued *N*-dimensional vector process of prices of traded securities with component processes $S_1(t),\ldots, S_N(t)$, which are strictly positive, adapted and right continuous with left limits (RCLL). Specify the first asset as a locally riskless money market account. By requiring $S_1(t)$ to have finite variation and to be continuous, one can write[4]

$$S_1(t) = exp\left(\int_0^t r(s)ds\right) \qquad 0 \leq t \leq T \tag{2.22}$$

for some process *r(t)*, which serves as the riskless interest rate at time *t*.

A portfolio is defined as an N-dimensional vector process $N=\{N(t), 0 \leq t \leq T\}$ with predictable components $N_1(t),\ldots,N_N(t)$ representing the quantities of the different assets in the portfolio and

$$\left(\int_0^t (N(s))^2 d[S]_s\right)^{\frac{1}{2}} \tag{2.23}$$

locally integrable. The value of the portfolio or trading strategy can again be expressed as a stochastic process $V= \{V_t(N(t)), 0 \leq t \leq T\}$ with initial value of

$$V_0(N(0)) = S(0)^T N(0) \tag{2.24}$$

and value at time *t* of

$$V_t(N(t)) = S(t)^T N(t) \tag{2.25}$$

A trading strategy N is self-financing if

$$V_t(N(t)) = V_0(N(0)) + \int_0^t N(s)dS(s) \quad \forall t, 0 \leq t \leq T \tag{2.26}$$

where the stochastic integral is defined by continuously extending the Lebesque-Stieltjes integral of simple predictable processes. From this relation, it can be seen that continuous-time economies are essentially the limit of multi-period economies.[5] We denote $\mathcal{N}$ as the set of these self-financing trading strategies.

An arbitrage opportunity of the first type is defined as a portfolio N with an initial value of zero

$$V_0(N(0)) = 0 \tag{2.27}$$

and a value at time T that is non-negative and strictly positive, with positive probability, that is,

$$V_T(N(T)) \geq 0 \quad \text{with probability one} \tag{2.28}$$

$$V_T(N(T)) > 0 \quad \text{with strictly positive probability} \tag{2.29}$$

Similar to the multi-period economy, an arbitrage opportunity of the second type exists if there is a trading strategy N such that

$$V_0(N(0)) < 0 \tag{2.30}$$

and

$$V_T(N(T)) \geq 0 \quad \text{with probability one} \tag{2.31}$$

For the same reasons discussed in the discrete case, these two types of arbitrage include all possible arbitrage opportunities. The following lemma states that an arbitrage opportunity of the second type implies an arbitrage opportunity of the first type. Although the exclusion of first-type arbitrage is therefore sufficient, and one can even argue why attention has been paid to second-type arbitrage opportunities, this general treatment stresses the similarities between discrete- and continuous-time economies.

Lemma 2.12 *If there exists an arbitrage opportunity of the second type in the continuous-time economy, then there also exists an arbitrage opportunity of the first type.*

Proof Suppose N^1 denotes an arbitrage opportunity of the second type in the continuous-time economy with initial value $V_0(N^1(0)) < 0$. *By* investing

$$N_1^2(0) := N_1^1(0) - V_0(N^1(0))$$

in the locally riskless money market account and

$$N_i^2(0) := N_i^1(0) \qquad \forall i = 2, \ldots, N$$

in the other traded securities, N^2 is an arbitrage opportunity of the first type in the continuous-time economy.■

Again, prices of traded securities are expressed in terms of the first asset by defining a discounted price process $S^*=\{S^*(t), 0 \leq t \leq T\}$ as

$$S_i^*(t) := exp\left(-\int_0^t r(s)ds\right)S_i(t) \qquad \forall i = 1, \ldots, N \tag{2.32}$$

Similarly, $V^* = \{V_t^*(N(t)), 0 \leq t \leq T\}$ denotes the discounted value process of the self-financing trading strategy[6] N

$$V_t^*(N(t)) = V_0^*(N(0)) + \int_0^t N(s)dS^*(s) \tag{2.33}$$

This transformation facilitates only the derivation of conditions under which arbitrage opportunities do not exist. The following lemma shows that no opportunities have been added or withdrawn because of this transformation.

Lemma 2.13 *There exists an arbitrage opportunity of the first type in the continuous-time economy if and only if there exists an arbitrage opportunity of the first type in the corresponding continuous-time economy, in which prices are expressed in terms of the value of the locally riskless money market account.*

Proof Suppose N is a trading strategy creating an arbitrage opportunity of the first type in the continuous-time economy. Because $S^*(0)=S(0)$,

$$V^*(N(0)) = 0$$

The prices of traded assets are strictly positive with probability one, and so the final value of this portfolio in the relative continuous-time economy is

$$V^*(N(T)) = \frac{1}{S_1(T)}S(T)^T N(T) \geq 0 \quad \text{with probability one}$$

$$V^*(N(T)) = \frac{1}{S_1(T)}S(T)^T N(T) > 0 \quad \text{with strictly positive probability}$$

The proof of the converse statement is almost similar.■

Thus, it has been shown that if an arbitrage opportunity in the continuoustime economy exists, an arbitrage opportunity of the first type exists in the relative continuous-time economy. As a logical consequence, conditions prohibiting this type of arbitrage in the relative economy also exclude any kind of arbitrage in the original continuous-time economy.

Theorem 2.14 *Arbitrage opportunities of the first type in the relative continuous-time economy do not exist if and only if there exists a probability measure* Q*, equivalent to P,* *such that discounted values of self-financing trading strategies are martingales with respect to this measure.*

Proof For a discussion regarding the necessity of the existence of an equivalent martingale measure under which self-financing trading strategies are martingales, see Harrison and Pliska (1981, p. 339). For a detailed proof for the case of a specific continuous-time economy in which there is one stock with a price process that follows a geometric Brownian motion and a money market account with a price process that is deterministic, see Heath and Jarrow (1987).■

The similarity between the results derived for the case of the continuoustime economy and the multi-period economy is striking. Theorem 2.14 is the exact equivalent of Theorem 2.8: if there is a probability measure such that self-financing trading strategies are martingales, arbitrage opportunities do not exist and vice versa. As a direct consequence in the multiperiod economy, this condition also implies that prices of different traded securities are martingales and vice versa. As the following Corollary shows, the converse statement no longer holds in continuous-time econ-omies. If relative prices of securities are martingales, relative values of trading strategies are local martingales.[7]

Corollary 2.15 *If relative values of self-financing trading strategies are martingales with respect to the probability measure* Q*, relative prices are martingales with respect to this probability measure. However, if prices are martingales with respect to the probability measure* Q*, relative self-financing trading strategies are local martingales with respect to this measure.*

Proof To prove the first statement, a self-financing trading strategy will be created by holding a long position in one of the N-traded assets. For a proof of the second statement, see Harrison and Pliska (1981, p. 238).■

This Corollary and Theorem 2.14 imply that trading strategies $N \in \mathcal{N}$ have to fulfil certain conditions to ensure that under a probability measure Q such that relative prices are martingales, relative trading strategies are martingales too. The class of local martingales is simply too large and if these conditions are not imposed, arbitrage strategies are still possible.

A well-known example of such a strategy is the doubling strategy discussed in Harrison and Kreps (1979) and Dybvig (1980). In this case, the initial and final trade are zero and one, respectively. The short-term rate of interest is zero and investors can trade at times $t_n = 1 - \left(\frac{1}{2}\right)^n$, for all n=0, 1, 2, 3,…. At each of these dates, a bet on the flip of a fair coin is possible. At time t_0, now, an investor makes a bet of 1 on heads. The initial value V_0 of this strategy is therefore

$$V_0 = 0 \tag{2.34}$$

If it turns out at time t_1 that heads occurs, he invests the proceeds in the riskless money market account and stops betting. If he loses, however, he withdraws 1 from the money market account and bets 2 on heads at the same time. At time t_n, the investor bets 2^n on heads, if all previous bets are lost. In case he wins, the proceeds cover previous losses plus one and are invested in the money market account. At time 1, the value of the strategy is[8]

V_1=1 with probability one

(2.35)

It is obvious from this example that the value of a trading strategy is not a martingale, although the possible outcomes of flipping a fair coin and the value of a money market account are.

Another example of a trading strategy that is a local martingale and not a martingale is the so-called suicide strategy, described in Harrison and Pliska (1981, p. 250). As in the previous example, the initial and final trading dates are zero and one, respectively, and the trading dates are $t_n = 1 - \left(\frac{1}{2}\right)^n$, for all n=0, 1, 2, 3,… . Investors have the opportunity to invest in a money market account or stock with corresponding price processes $S_1=\{S_1(t), 0 \le t \le 1\}$ and $S_2=\{S_2(t), 0 \le t \le 1\}$, respectively. The initial value of the stock is one and without loss of generality, it can be assumed that interest rates are deterministic and zero, such that $S_1(t)=1$ for all $0 \le t \le 1$. At time 0, an investor sells b, $b > 0$ shares of stock short and puts the proceeds plus one in the money market account, that is,

$$N_1(0) = 1 + b$$

(2.36)

$$N_2(0) = -b$$

The initial value of this strategy therefore is

$$V_0(N(0)) = 1$$

(2.37)

The probability of ruin during the interval $[t_0, t_1]$ is $p=Pr(inf\{t: S_2(t) = 1+1/b\})$. If ruin actually occurs before t_1, the portfolio is liquidated. If not, the number of shares sold short is increased at time t_1 such that during the interval $[t_1, t_2]$, the probability of ruin is again equal to p. At time 1, then, the value of the portfolio is[9]

$V_1(N(1))=0$ with probability one

(2.38)

Because this strategy can be performed with any positive initial value, values of trading strategies that yield the same final value are not unique. Suppose there is a trading strategy N_1 with initial value $V_0(N_1(0))$ and final value $V_T(N1(T))$. Create another strategy from this, $N_2(t)= N_1(t)+N(t)$ for all $0 \le t \le T$, with a some positive real-valued constant. The final value of this strategy also is $V_T(N_1(T))$. However, the initial value equals $V_0(N_1(0))+\alpha$.

A number of authors have proposed certain trading constraints that are exogenously imposed upon the continuous-time economy. These constraints restrict the class of local martingales to martingales and therefore ensure that arbitrage opportunities are prohibited.

One possible solution, as proposed by Kreps (1979), is to impose a uniform instead of local bound on the predictable vector process N. Although the total number of securities outstanding would be a natural candidate and does not seem to restrict trading strategies, existing short sale positions are ignored and the uniform bound is inconsistent with frictionless security markets in which securities are infinitely divisible.

Another possible trading restriction is the specification at the initial date of a finite number of dates at which investors are allowed to trade. These so-called simple trading strategies are extensively discussed in Harrison and Kreps (1979), but are in a sense too restrictive. One cannot ensure that every contingent claim is attainable unless additional structure upon the preferences of investors is defined.[10]

Some constraint on the value of the investor's entire portfolio also serves as a trading restriction under which the relative value of a portfolio is a martingale if relative prices are martingales. This constraint is investigated by Dybvig (1980) (the case of general wealth constraints) and is explained in Heath and Jarrow (1987) (the specific case of margin requirements). Because these requirements are actually present in security markets and do not seem too restrictive as far as the attainability of contingent claims is concerned, it will be assumed throughout the rest of the thesis that this restriction holds.

Let denote the set of equivalent probability measures such that relative prices of traded securities are martingales. Given this set $\mathcal{Q}$, then, $\mathcal{N}', \mathcal{N}' \subset \mathcal{N}$ denotes the set of self-financing trading strategies such that the relative values of these portfolios are martingales, too. As in the multi-period economy, $\mathcal{X}$ is the set of contingent claims {Xt, $0 \leq t \leq T$} and it is assumed that the processes X_t relative to the money market account are integrable and adapted to the filtration $\mathcal{F}_t$. From the discussions in the previous section, this set is known to be general enough to contain all possible claims. Again, call a contingent claim X_t to be attainable if there exists a trading strategy $N \in \mathcal{N}'$ that generates the claim $Vt(N(t))=Xt$; call $\pi(s)=V_s(N(s)),\ 0\leq s\leq t$ *the* price of the contingent claim at time *s*. Before the last theorem regarding the attainability of contingent claims is presented, the following theorem ensures that if a contingent claim is attainable and arbitrage opportunities are prohibited, the price of the claim is unique.

Theorem 2.16 *If arbitrage opportunities do not exist in the continuoustime economy, there is a single price π(s) associated with any attainable con*tingent claim $X_t \in \mathcal{X},\ 0 \leq s \leq t \leq T$, which satisfies $\pi(s) = E_Q\left(\frac{S_1(s)}{S_1(t)} X_t | \mathcal{F}_s\right)$.

Proof See the proof of Corollary 2.9.■

The following theorem now states that if the set of equivalent martingale measures $\mathcal{Q}$ is a singleton, every contingent claim is attainable.

Theorem 2.17 *The continuous-time economy contains no arbitrage opportunities and is complete if there exists a unique equivalent probability measure such that relative self-financing trading strategies are martingales.*

Proof See Harrison and Pliska (1981, Corollary 3.36, p. 241).■

The Black and Scholes (1973) formula will now be derived by closely following the lines suggested by the theorems above. This example serves as a nice illustration of the strong implications of the exclusion of arbitrage opportunities and the valuation of contingent claims and illustrates, more generally, the procedure to be followed in the next chapters when interest rate contingent claims will be discussed.

The continuous-time economy in which Black and Scholes derived a closed-form solution for a European call option on a non-dividend paying stock is mainly characterized by a Brownian motion $\{W(t), 0 \le t \le T\}$ defined on a filtered probability space $(\Omega, \mathcal{F}, F, P)$ described at the beginning of this section. We focus on the stochastic process of one stock price, $\{S(t), 0 \le t \le T\}$, which follows a geometric Brownian motion and pays no dividends,

$$S(t) = S(0) + \int_0^t \mu S(s)ds + \int_0^t \sigma S(s)dW(s) \tag{2.39}$$

Because S(0), μ and σ are assumed to be strictly positive constants, the stochastic differential equation above has a unique solution.[11] The riskless interest rate r is assumed to be constant, resulting in the following deterministic differential equation for the money market account $\{B(t), 0 \le t \le T\}$,

$$B(t) = 1 + \int_0^t rB(s)ds \tag{2.40}$$

The first step in the valuation procedure is the derivation of the stochastic differential equation of the relative value of the stock price $\{S^*(t)=S(t)/B(t), 0 \le t \le T\}$. Because the function $f(x_1, x_2)=x_1/x_2$, $x1, x2 > 0$ is continuously differentiable, Ito's Lemma can be applied to obtain the unique process of $S^*(t)$

$$S^*(t) = S^*(0) + \int_0^t (\mu - r)S^*(s)ds + \int_0^t \sigma S^*(s)dW(s) \tag{2.41}$$

The next step is to show that there exists a unique equivalent martingale measure such that the relative price $S^*(t)$ is a martingale. From Theorem 2.17 it is known that arbitrage opportunities do not exist and

that every contingent claim is attainable. To show that is actually the case, first A. is defined as the market price of risk,

$$\lambda = \frac{\mu - r}{\sigma} \tag{2.42}$$

and the Radon-Nikodym derivative {p(λ, t), 0 /), $0 \le t \le T$}, which defines the change of probability measure, is specified as follows

$$\rho(\lambda, t) = exp\left(\int_0^t \lambda dW(s) - \frac{1}{2}\int_0^t \lambda^2 ds\right) \tag{2.43}$$

Due to the assumptions regarding the coefficients of the stochastic process for S(t) and r, the market price of risk is a constant, implying

$$E_P(\rho(\lambda, t)|\mathcal{F}_0) = 1 \qquad 0 \leq t \leq T \tag{2.44}$$

It is now allowed to apply Girsanov's Theorem,[12] which states that the process $\{\tilde{W}(t), 0 \leq t \leq T\}$ defined by

$$\tilde{W}(t) = W(t) + \int_0^t \lambda ds \tag{2.45}$$

is a Brownian motion for the filtered probability space $(\Omega, \mathcal{F}, F, Q)$. The unique equivalent probability measure Q is given by

$$dQ = \rho(\lambda, T)dP \tag{2.46}$$

By the linearity of stochastic integration, and equations (2.41), (2.42) and (2.45), one can write

$$S^*(t) = S^*(0) + \int_0^t \sigma S^*(s) d\tilde{W}(s) \tag{2.47}$$

which is a martingale with respect to the probability measure Q[13] Given the information set at some time *s*, *S*(t)*, $t \geq s$ is lognormally distributed[14] with mean *S*(s)* and variance *S*(s)2(exp(σ2(t–s))*–1).

The last step is the actual calculation of the value of a European call option maturing at time $0 < \tau \leq T$ with exercise price $K > 0$. The value of the call at maturity is given by

$$C(S(\tau), K, \tau) = Max(S(\tau) - K, 0) \tag{2.48}$$

To obtain the value of the call at some time *t*, $0 \leq t \leq \tau$, apply Theorem 2.16

$$C(S(t), K, t) = E_Q\left(\frac{B(t)}{B(\tau)} C(S(\tau), K, \tau) \mid \mathcal{F}_t\right) \tag{2.49}$$

which is equal to

$$C(S(t), K, t) = B(t) E_Q\left(Max\left(S^*(\tau) - \frac{K}{B(\tau)}, 0\right) \mid \mathcal{F}_t\right) \tag{2.50}$$

or, equivalently,

$$C(S(t),K,t) = B(t)E_Q\left(S^*(\tau) \mid \mathcal{F}_t \cup \left\{S^*(\tau) > \frac{K}{B(\tau)}\right\}\right) - \tag{2.51}$$

$$K\frac{B(t)}{B(\tau)}Pr_Q\left(S^*(\tau) > \frac{K}{B(\tau)} \mid \mathcal{F}_t\right)$$

Similar to the example at the end of the previous section, $Pr_Q\left(S^*(\tau) > \frac{K}{B(\tau)} \mid \mathcal{F}_t\right)$ denotes the probability under the measure Q of the event $S^*(\tau) > \frac{K}{B(\tau)}$ given the information at time *t.* After some straight-forward calculations,

$$C(S(t),K,t) = S(t)N(d_1) - K\frac{B(t)}{B(\tau)}N(d_2) \tag{2.52}$$

where

$$d_1 = \frac{log\left(\frac{S(t)B(t)}{KB(\tau)}\right) + \frac{1}{2}\sigma^2(\tau - t)}{\sigma\sqrt{\tau - t}}$$

$$d_2 = d_1 - \sigma\sqrt{\tau - t}$$

and *N(.)* is the cumulative standard normal distribution function.

From this closed-form formula, it is evident that the value of the contingent claim at time *t*, $0 \le t \le \tau$ is equal to a portfolio containing the stock *S(t)* and the money market account *B(t)* with weights $N(d_1)$ and $\frac{K}{B(\tau)}N(d_2)$, respectively. It is easy to verify that this trading strategy is self-financing and that the final value of the portfolio is equal to the value of the call option at maturity.

Another interesting issue is the fact that the call option formula is independent on the coefficient μ. As has been shown in the previous discussion, investors have equivalent probability measures if they agree on the value of σ, whatever their individual assessment of μ is. Because the exclusion of arbitrage opportunities and the attainability of contingent claims is guaranteed if there exists a unique equivalent probability measure such that relative prices are martingales, this independence is easily explained.

NOTES

1 For more on feasibility and the relationship between primal and dual linear programming problems, see Chvátal (1983).

2 Although this theorem can be proved straightforwardly and in a sense equivalently by using Farkas' Lemma or the Separating Hyperplane Theorem (see, for example, Pedersen *et al.* 1989) this approach is preferred here because of its clear interpretation.

3 For an intuitive explanation of probabililty spaces in continuous-time econ-omies, see Duffie (1988, ch. 14).
4 See Harrison and Pliska (1981, p. 232).
5 See Duffie (1988, ch. 15).
6 Again, see Harrison and Pliska (1981, p. 238) for a formal proof of the equivalence between a self-financing portfolio *N* in the continuous-time economy and in the relative continuous-time economy.
7 For a definition of local martingales, see Harrison and Pliska (1981, p. 235).
8 Because

$$Pr(\text{Winning 1 at time 1}) = 1 - Pr(\text{Losing all bets at time 1}) = 1 - \lim_{n\to\infty}\left(\frac{1}{2}\right)^n = 1$$

9 Because

$$Pr(\text{Ruined at time 1}) = 1 - Pr(\text{Survived at time 1}) = 1 - \lim_{n\to\infty}(1-p)^n = 1$$

10 Brennan (1979) showed that under the assumption of a bivariate lognormal distribution of the price of the underlying asset and aggregate wealth, a sufficient condition to obtain the Black-Scholes formula is the utility function to exhibit constant proportional risk aversion. Duan (1990) extended this valuation technique in the case of *Garch(p, q)* distributed lognormal returns.
11 See Gihman and Skorohod (1972, p. 40).
12 For the details of Girsanov’s Theorem, see Elliott (1982).
13 See, for example, Duffie (1988, p. 142).
14 For a clear explanation of the distribution of linear stochastic differential equations, see Arnold (1974, ch. 8).

3
AN OVERVIEW OF THE VALUATION OF INTEREST RATE DERIVATIVE SECURITIES

The previous chapter examined the notion of arbitrage opportunities and the conditions under which these opportunities are excluded in security markets. The valuation of derivative securities, and especially interest rate derivative securities, is strongly related to this concept and the accompanying conditions. It was shown that every contingent claim can be replicated by a portfolio, consisting of traded securities, if there exists a unique equivalent probability measure such that the value of trading strategies expressed in terms of a money market account is a martingale. This strong result actually makes it possible to obtain the value of any contingent claim by calculating the expected discounted value of the claim at maturity. This expectation has to be taken with respect to this unique probability measure.

This chapter serves as an introduction to the next chapters, which will thoroughly discuss the various models developed to value interest rate derivative securities. First, a general overview of the different valuation approaches, using Figure 3.1, will provide the reader with some feeling for valuing interest rate derivative securities. The first section of this chap-ter, therefore, discusses the valuation of interest rate derivative securities based on the explicit modelling of the underlying values. This approach, which is called the direct approach, is similar to the stock option valuation problem, solved in the seminal paper of Black and Scholes (1973). The second section then illustrates the indirect approach, in which the values or prices of all interest rate dependent securities are considered to be a function of the instantaneous short rate. Because the main objective of this chapter is the general comparison of the characteristics of each of the valuation approaches, not so much attention will be given to exact definitions of probability spaces and regularity conditions of stochastic processes.

THE DIRECT APPROACH

The first step in valuing derivative securities in the direct approach is the specification of the stochastic behavior of the underlying values. Given

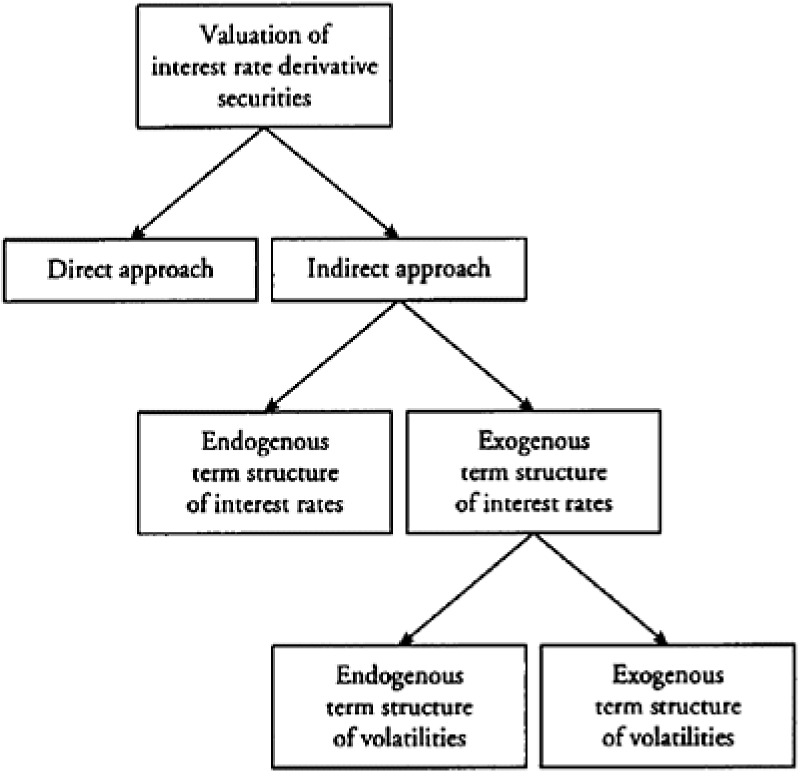

Figure 3.1 Overview

these specified characteristics of the underlying values and the properties of the contingent claim, the exclusion of arbitrage opportunities and the desire for a unique price of this claim require the existence of an equivalent probability measure as discussed in the previous chapter. A nice example of this approach is the Black and Scholes (1973) European call option formula, which has been derived in the previous section. An explanation of the direct approach starts with the general specification of a continuous time vector process $B=\{B(t, \theta 1), 0 \leq t \leq \tau\}$, denoting the prices of the underlying values of the derivative security. The initial and final trade date are 0 and τ, respectively, and the symbol θ_1 represents the specific characteristics of the underlying securities, such as time-to-maturity and coupon payments. Before deriving the unique value of the contingent claim, define $C=\{C(t, \theta 2), 0 \leq t \leq \tau\}$ as the stochastic process of the claim. At the maturity date, the claim is a known function of the values of the underlying securities during maturity. The characteristics of this claim, such as exercise prices, are represented by the symbol θ_2 and it is assumed that the final trading date τ equals the maturity of the derivative security. This may seem rather restrictive. Because the main purpose of the direct approach is the valuation of the contingent claim, however, only the values of this claim and its underlying securities during maturity of the claim are of interest.

It will be further assumed that one of the securities, say $B_1(t, \theta_1)$, can be regarded as an alternative default-free investment, maturing at the final trade date τ. In the previous chapter, this alternative investment was a locally riskless money market account that allowed investors to invest or withdraw money at an instantaneously riskless interest rate. The existence of a unique probability measure such that relative trading strategies are martingales, then also implied the opportunity to create a strategy that replicates the value of a default discount bond maturing at τ. In the direct approach, however, only this latter opportunity has to exist. As already mentioned above, because the principal objective is the unique arbitrage-free valuation of one particular contingent claim, arbitrage opportunities between this claim, its underlying values and the alternative investment have to be excluded. As such, this approach can also be regarded as a partial equilibrium approach because no conditions are imposed to ensure that arbitrage opportunities between other contingent claims and other traded securities are also prohibited.

This section started with a general specification of the stochastic process of the underlying traded securities. No reference has been made with respect to the probability space and the corresponding probability measure. To be able to obtain a unique arbitrage-free value of the contingent claim, however, it must be assumed that there exists a unique probability measure Q such that the values of trading strategies consisting of the underlying values expressed in terms of the alternative investment are martingales. In that case, arbitrage opportunities between these strategies and the alternative investment are prohibited and the unique value of the contingent claim C at time t, $0 \leq t \leq \tau$ equals

$$C(t, \theta_2) = B_1(t, \theta_1) E_Q(C(\tau, \theta_2)) \qquad 0 \leq t \leq \tau \tag{3.1}$$

In this expression, $E_Q(.)$ denotes the expectation at time t with respect to the probability measure Q and it is assumed that the claim is a suitable regular function of the underlying values such that this expectation exists.

This explanation of the basic steps and characteristics of the direct approach sets the stage for a discussion of some of the advantages and limitations. For a European call option price on a coupon paying bond, both the stochastic processes of the underlying bond and a discount bond with a maturity equal to the option must be specified. As a realistic description of the actual behavior of bond prices, these price processes must incorporate the payment of coupons and ensure that at maturity of the bond the value of the bond equals its face value. In addition to these obvious requirements, nominal yields of bonds are known to be always positive, which puts another constraint on the specification of the processes.

After this specification, the value of the option can be determined by actually calculating the discounted expectation of the option at maturity. The expectation has to be taken with respect to the unique probability measure such that the value of the coupon paying bond in terms of the discount bond is a martingale. As is clear from the example of the European call option on the coupon paying bond, the derivation of the existence of an equivalent martingale measure and the determination of the option's price under this measure can be difficult.

The basic advantage of this approach, however, is the fact that only the processes of those securities on which the claim is dependent have to be specified. No assumptions have to be made, therefore, about a general equilibrium within the fixed-income market. As another result, once the processes of specific bonds are obtained, the observed prices of these bonds can be used to estimate the necessary parameters. As will be seen in the next section, this is rarely the case in the indirect approach.

THE INDIRECT APPROACH

The main characteristic of the direct approach is the explicit formulation of the stochastic processes of the securities on which the contingent claim is dependent. As discussed in the previous section, a severe disadvantage of this approach is the fact that these processes have to capture specific features of the security, such as a known face value at the maturity of a bond. Fulfilling these requirements and ending up with processes that allow for the calculation of the value of the contingent claim has proven difficult.

The indirect approach, however, starts with the specification of some processes on which all interest rate dependent securities depend. Generally, the first step is the assumption that the process of the instantaneous spot rate $\{r(t), 0 \leq t \leq \tau\}$ is a well-defined function of the values of these basic processes. In the next step then, all interest rate dependent securities, represented by the vector process $B=\{B(t, \theta_1), 0 \leq t \leq \tau\}$, are considered to be a function of this instantaneous spot rate. In this formulation, 0 and τ are the initial and final trade date, respectively, and θ_1 denotes the characteristics of the securities. Because the underlying assets of the contingent claim are in fact derivative securities, too, the unique arbitrage-free prices of these securities must also be derived. Without a detailed description of the probability space, it will be assumed for the moment that sufficient conditions are fulfilled to ensure that these unique prices exist. From the previous chapter, this means that there exists a unique probability measure Q such that values of trading strategies relative to the money market account are martingales, or equivalently,

$$B(t,\theta_1) = E_Q\left(e^{-\int_t^{t'} r(s)ds} B(t',\theta_1)\right) \qquad 0 \leq t \leq t' \leq \tau \tag{3.2}$$

It is obvious that E_Q, again, denotes the expectation operator under the measure Q.

The last step is the actual calculation of the value of the contingent claim. Define $C=\{C(t, \theta 2), 0 \leq t \leq \tau\}$ as the stochastic process of the claim and let the symbol $\theta 2$ denote the characteristics of the contingent claim. At maturity of the claim, say τ_1, $0 \leq \tau_1 \leq \tau$, the payout of the claim is a known suitable regular function of the values of the underlying securities during maturity. As with the values of the underlying securities, the unique arbitrage-free price of the derivative security can be obtained by calculating the discounted expectation with respect to the martingale measure Q, or

$$C(t,\theta_2) = E_Q\left(e^{-\int_t^{t'} r(s)ds} C(t',\theta_2)\right) \qquad 0 \leq t \leq t' \leq \tau_1 \leq \tau \tag{3.3}$$

Because all interest rate dependent securities are assumed to be functions of the instantaneous spot rate, the indirect approach can also be considered as a general equilibruim approach. Given the stochastic process of this short rate, the prices of bonds with various maturities or the term structure of interest rates at some valuation date can be derived. After this valuation date, the process of the short rates implies a stochastic behavior of the term structure of spot rates, or, which is equivalent, a term structure of spot rate volatilities. The various general equilibrium models developed in past years can now be simply classified according to these two term structures.

The first distinction can be made with respect to the term structure of interest rates. If the process of the instantaneous spot rate is explicitly modelled in the first step as a function of some parameters, the term structure of spot rates at the valuation date is a specific function of these parameters. The possible shape of this term structure, therefore, is endogenously implied by the stochastic characteristics of the short rate. As will be shown in later chapters, where some of these models are tested empirically, this relationship can also be used the other way around. In that case, the shape of the term structure of interest rates is considered to imply a stochastic behavior of the instantaneous spot rate.

The class of exogenous term structure of interest rates models, however, starts at the valuation date with a given term structure of interest rates. The parameters of the stochastic process of the short rate are time-dependent to ensure that this observed term structure is obtained at the valuation date. It is obvious that in this class of models no inferences can be made at the valuation date about the probability distribution of the instantaneous spot rate.

The second distinction, which is only relevant in the case of the exogenous term structure of interest rate models, can be made with respect to the term structure of interest rate volatilities. The time-dependent parameters of the instantaneous spot rate not only ensure a pre-specified term structure of spot rates at the valuation date, but also imply a particular shape of the term structure of interest rate volatilities. This term structure can have a specific functional form, which is dependent on a few parameters. These parameters, however, can also be taken in such a way that a given volatility structure of spot rates is implied. As before, the same distinction now applies.

The basic difference between the direct and indirect approaches is the modelling or specification of the securities on which a contingent claim is dependent. A serious limitation of the direct approach is the formulation of stochastic processes describing the stochastic behavior of the underlying securities. As discussed in the previous section, this can be rather difficult in the case of coupon paying bonds because of numerous boundary conditions. In the indirect approach, however, this problem is to some extent avoided. In the first step, a stochastic process describing the stochastic behavior of the instantaneous spot rate must be formulated. In the second step, then, the stochastic process of the underlying securities is derived by taking the discounted expectation of the payouts. The value of the coupon paying bond, for example, is simply the sum of the values of the different payments. The specific features of securities are thus modelled indirectly in the second step and do not have to be modelled explicitly in the first step, as in the direct approach.

Another difference between the two approaches is the resulting impact on a possible equilibrium within the fixed-income market. As all interest dependent securities are solely dependent on the instantaneous spot rate, it has been seen that the indirect approach requires a general equilibrium between these securities. On the one hand, this can be regarded as a strong result. The general modelling of the term structure of interest rates allows for the valuation of interest rate derivative securities, for which the underlying value is not necessarily observable. An example of such a security is a callable bond, which can be seen as an ordinary coupon paying bond less the value of a call option on this bond. To value this bond in the direct approach, a specification of the price process of the underlying bond is needed even if this bond does not exist. On the other hand, the description of the term structure and its likely movement over time implies a stochastic behavior of specific securities. As will be seen, it does not need to be true that this particular behavior is suitably described by the general stochastic movement of the term structure.

4
MODELLING BOND PRICES

The valuation of European call and put options on stocks is one of the most popular examples in financial research of the strong implications of the no-arbitrage approach. Given a Geometric Brownian Motion to describe the stochastic behavior of stocks over time and a constant interest rate at which investors can finance their hedge positions, Black and Scholes (1973) were the first to derive a closed-form solution for the value of a European call and put option on a stock. This approach was extended in a seminal paper by Merton (1973) to incorporate a stochastic instead of a constant short-term rate of interest.

In utilizing this approach to the valuation of options on discount bonds, some fundamental differences are encountered between the possible stochastic behavior of stocks and bonds that affect the value of an option. At any time before maturity of the bond, for example, the investor knows with certainty that at maturity he will receive the principal payment of the bond. As a result, the stochastic process that describes the bond price has to ensure that at maturity the value of the bond equals this face value. In addition, the uncertainty of bond prices or the variance of the corresponding bond returns is decreasing during the maturity time of the bond and is zero at maturity. Finally, because of the one-to-one correspondence between bond prices and interest rates, it is clearly inconsistent to assume the short-term rate of interest to be constant as in the original derivation of Black and Scholes. Although one can argue that the value of short-term options on long-term bonds is hardly affected by these fundamental differences, a generally applicable bond option model should incorporate these specific bond characteristics.[1]

The model of Schaefer and Schwartz (1987) modifies the stochastic differential equation of the bond price by assuming that the volatility is proportional to the duration of the bond. As maturity decreases, the volatility decreases and becomes zero in the end. However, the drift term of the Schaefer and Schwartz bond price process is assumed to be constant, ignoring thereby the above-mentioned price effect, by which bond prices are forced to equal their face value at the final maturity date.

In addition, the short-term rate of interest is assumed to be constant, restricting the empirical applicability of their model to short-term options on long-term bonds.

The model of Ball and Torous (1983) incorporates the “drift-to-face-value” effect by using the Brownian Bridge to model bond prices. A significant weakness of their model, however, is the assumed constant instantaneous variance of bond prices. The volatility of the corresponding yield-to-maturity, therefore, increases without bound as the bond approaches maturity, which again stresses the importance of consistently modelling the different specific bond price features. To obtain arbitrage-free option values, as Chapter 2 pointed out, a unique equivalent probability measure such that relative bond prices are martingales has to exist. Although Ball and Torous actually derive closed-form solutions for Euro-pean options on discount bonds by simply assuming such an equivalent martingale measure to exist,[2] Cheng (1989) has shown that this measure does not exist and that arbitrage opportunities are, therefore, not excluded. Referring to the above-mentioned volatility of the yield-to-maturity implied by their bond price process, this result is hardly surprising.

This chapter develops a model incorporating both the volatility specification of the Schaefer and Schwartz model and the “drift-to-face-value” effect of the model of Ball and Torous. Starting with a specification of the stochastic process of the yield-to-maturity of a bond with constant coefficients, the corresponding stochastic process of the bond prices is easily obtained. The instantaneous volatility of the resulting stochastic differential equation is linearly dependent on the remaining time-to-maturity or duration of the discount bond. In addition, the drift term of the process is equal to that found in the Ball and Torous model, ensuring that at maturity the value of the bond equals its face value. Although it is argued in Cheng (1989, p. 196) that, in general, a Brownian Bridge process to model bond prices is not acceptable to value options (as the equivalent martingale measure does not exist), it will be shown in the current model that the modified Brownian Bridge process does allow for a unique change of measure such that relative prices are martingales. Based on this equivalent measure then, closed-form solutions will be obtained for the value of European call and put options on discount bonds.

Apart from the theoretical validity of the Schaefer and Schwartz model, the valuation of options is exactly the same as that found in Merton (1973). Although the drift term of bond prices is irrelevant for the bond option price, a result which is thoroughly discussed at the end of Chapter 2 in the case of the Black and Scholes model, the estimation of the volatility parameters is definitely dependent on the particular functional form of the drift term. The estimation procedure suggested by Ball and Torous, however, ignores this time-varying drift term, although their model gets its strength from a theoretical drift term, which.is not constant. This chapter obtains maximum likelihood estimators for the volatility parameters of the option price that are consistent, asymptotically normally distributed, and efficient in the class of all consistent and uniformly asymptotically normal (CUAN) estimators. The derived theoretical estimators for the different parameters are also easily applicable to a sample of observations where the length of the time intervals between subsequent observations is not constant, due to non-trading.

The results of this chapter with respect to the theoretical validity of the discount bond option model and the estimation of the necessary input parameters are supposed to enhance the understanding of the valuation of interest rate derivative securities and may improve current research. A practical implementation of this approach is not yet possible because no options except those on coupon paying bonds are traded and a natural extension of the proposed model to incorporate these additional price characteristics is not straightforward, as will be made clear in the next sections.

The first section of this chapter presents the current model and shows that the stochastic differential equations of the bond price have a unique solution. The second section demonstrates that a unique equivalent martingale measure exists, thereby excluding any arbitrage opportunities. After this, closed-form solutions are derived for the values of European call and put options on a discount bond. Finally, the third section derives the maximum likelihood estimators of the different coefficients or parameters of the bond price process necessary for the valuation of options on bonds in the second section.

THE MODEL

The continuous-time economy is characterized by an initial trade date 0 and a final trade date T_S. The continuous-time uncertainty is specified by the filtered probability space $(\Omega, \mathcal{F}, F, P)$. Ω denotes the state space, P some probability measure and F the filtration of increasing σ-algebras $\mathcal{F}_t$, $0 \leq t \leq T_S$, which satisfy the usual conditions. As before, investors have only to agree on the null sets of the probability measure, which implies that the measure P may be replaced by an equivalent probability measure P^*.

Suppose $P_L(t, T_L)$ denotes the value at time t [0, Ts] of a discount bond that matures at time $T_L > T_S$ and has unit face value. Suppose, further, that $Ps(t, T_S)$ is the value of a similar bond maturing at time Ts. Given the value of these discount bonds at some time, the corresponding yield-to-maturity is obtained from the following familiar relationship:

$$P_L(t, T_L) = e^{-y_L(T_L - t)} \qquad (4.1)$$

$$P_S(t, T_S) = e^{-y_S(T_S - t)}$$

The unique Ito processes of these yield-to-maturities $(y_L, y_S) = \{(y_L(t), y_S(t)).\ t$ [0, T_S]} obey the following stochastic differential equations,

$$\begin{pmatrix} dy_L(t) \\ dy_S(t) \end{pmatrix} = \begin{pmatrix} \alpha_L \\ \alpha_S \end{pmatrix} dt + \begin{pmatrix} \sigma_{L_1}\sigma_{L_2} \\ \sigma_{S_1}\sigma_{S_2} \end{pmatrix} \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix} \qquad (4.2)$$

with

$$\begin{pmatrix} y_L(0) \\ y_S(0) \end{pmatrix} \in \mathbb{R}^2;\ \begin{pmatrix} \alpha_L \\ \alpha_S \end{pmatrix} \in \mathbb{R}^2;\ \begin{pmatrix} \sigma_{L_1}\sigma_{L_2} \\ \sigma_{S_1}\sigma_{S_2} \end{pmatrix} \in \mathbb{R}^2 \times \mathbb{R}^2$$

The two-dimensional process $W = \{(W_1(t), W_2(t), t$ [0, T_S]} is a vector process of independent Standard Brownian Motions initialized at zero and defined on the above-specified probability space $(\Omega, \mathcal{F}, F, P)$

The relationship between the discount bonds and their respective yields allows the determination of the unique stochastic Ito processes $P_L = \{P_L(t, T_L), t$ [0, T_S]} and $P_S = \{P_S(t, T_S), t$ [0, T_S]} by using Ito's Lemma,[3]

$$\begin{pmatrix} dP_L(t, T_L) \\ dP_S(t, T_S) \end{pmatrix} = \begin{pmatrix} \mu_L(t) P_L(t, T_L) \\ \mu_S(t) P_S(t, T_S) \end{pmatrix} dt - \begin{pmatrix} P_L(t, T_L) & 0 \\ 0 & P_S(t, T_S) \end{pmatrix} \times \qquad (4.3)$$

$$\Omega(t) \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix}$$

with

$$\mu_L(t) = -\frac{\ln P_L(t, T_L)}{T_L - t} + \frac{1}{2}(\sigma_{L_1}^2 + \sigma_{L_2}^2)(T_L - t)^2 - \alpha_L(T_L - t)$$

$$\mu_S(t) = -\frac{\ln P_S(t, T_S)}{T_S - t} + \frac{1}{2}(\sigma_{S_1}^2 + \sigma_{S_2}^2)(T_S - t)^2 - \alpha_S(T_S - t)$$

and

$$\Omega(t) = \begin{pmatrix} \sigma_{L_1}(T_L - t) & \sigma_{L_2}(T_L - t) \\ \sigma_{S_1}(T_S - t) & \sigma_{S_2}(T_S - t) \end{pmatrix}$$

The following theorem asserts that the process given by equation (4.3) actually fulfils suitable regularity conditions that guarantee the existence of a unique solution.[4]

Theorem 4.1 *There exists a unique, continuous and square integrable stochastic process* $(P_L, P_S)=\{(P_L(t, T_L), P_S(t, T_S)), t\ [0, T_S]\}$ *that is a solution to equation (4.3).*

Proof See Appendix A.■

It is interesting now to investigate exactly the differences between the Brownian Bridge model described above and the one proposed by Ball and Torous (1983). In their model, the stochastic process of the short-term and long-term bonds $(P_L, P_S)=\{(P_L(t, T_L), P_S(t, T_S)), t\ [0, T_S]\}$ obey the following stochastic differential equation,[5]

$$\begin{pmatrix} dP_L(t, T_L) \\ dP_S(t, T_S) \end{pmatrix} = \begin{pmatrix} \mu_L^{BT}(t)P_L(t, T_L) \\ \mu_S^{BT}(t)P_S(t, T_S) \end{pmatrix} dt - \begin{pmatrix} P_L(t, T_L) & 0 \\ 0 & P_S(t, T_S) \end{pmatrix} \times$$

$$\Omega^{BT}(t)\begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix}$$

with drift term

$$\mu_L^{BT}(t) = -\frac{\ln P_L(t, T_L)}{T_L - t} + \frac{1}{2}(\sigma_{L_1}^2 + \sigma_{L_2}^2)$$

$$\mu_S^{BT}(t) = -\frac{\ln P_S(t, T_S)}{T_S - t} + \frac{1}{2}(\sigma_{S_1}^2 + \sigma_{S_2}^2)$$

and covariance matrix

$$\Omega^{BT}(t) = \begin{pmatrix} \sigma_{L_1} & \sigma_{L_2} \\ \sigma_{S_1} & \sigma_{S_2} \end{pmatrix}$$

The difference between this expression and the model given by (4.3) is the multiplication of the volatility parameters by the remaining time-to-maturity of the bond or, which is equivalent, the duration of the bond. As already noted in the introduction to this chapter, the “drift-to-face-value” of the bond implies a decreasing instantaneous variance that reaches zero at maturity. To explain the importance of this implication, the following stochastic process of the corresponding yield-to-maturities of the Brownian Bridge model of Ball and Torous is derived, which is equal to

$$\begin{pmatrix} dy_L(t) \\ dy_S(t) \end{pmatrix} = \begin{pmatrix} \frac{\sigma_{L_1}}{T_L - t} & \frac{\sigma_{S_1}}{T_L - t} \\ \frac{\sigma_{L_2}}{T_S - t} & \frac{\sigma_{S_2}}{T_S - t} \end{pmatrix} \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix} \tag{4.5}$$

Although the above short derivation is not formally justified because the yield processes are not Ito processes and the initial bond price processes do not fulfil sufficient conditions guaranteeing uniqueness and existence, it does give the necessary economic intuition as to why their model is not arbitrage-free and why the variance of the bond price should be a decreasing function of its remaining time-to-maturity.

Another interesting implication of the Brownian Bridge process of Ball and Torous is the fact that negative interest rates occur with probability one. As noted by Ball and Torous (1983, p. 524), a Brownian Bridge process $Z(t), t$ [0, 1] can be seen as the following transformation of a Standard Brownian Motion $W(t), t$ [0, 1],

$$Z(t) = (1 - t)W\left(\frac{t}{1 - t}\right) \tag{4.6}$$

The function $t \rightarrow \frac{t}{1-t}$ gives a one-to-one correspondence between the intervals (0, 1) and (0, ∞). Because it is well known that a Brownian Motion, independent of its particular state, will become zero at some future point in time with probability one, the Brownian Bridge becomes zero at some time t [0, 1]. In the current model, the effective probability of negative interest rates is much smaller.[6]

OPTION VALUATION

The possibility of the arbitrage-free valuation of a claim contingent on some security is equivalent to the existence of a unique equivalent probability measure such that the relative price of the security is a martingale. Chapter 2 showed that if such a probability measure indeed exists, the value of any claim should be equal to its discounted expected value under the equivalent martingale measure.

The previous section presented a description of the stochastic evolution of bond prices over time. It was explained that the instantaneous variance of bond prices should be a decreasing function of a bond's remaining time-to-maturity or duration. This particular property appears to be the crucial difference between the current model and the Brownian Bridge model as proposed by Ball and Torous. Correspondingly, as shown by Cheng (1989), their model does not allow for the arbitrage-free valuation of contingent claims because the above-mentioned unique probability measure does not exist.

This section will show that in the current model there does exist a unique probability measure such that the stochastic process of the longterm maturity bond $P_L(t, T_L)$ in terms of the value of the short-term maturity bond $P_S(t, T_S)$ is a martingale. After this, the exact values will be derived of European call and put options with maturity T_S written on the discount bond $P_L(t, T_L)$, and some of the properties of these valuation formulas will be discussed.

Because the stochastic differential equation of the different processes $P_L=\{P_L(t, T_L), t$ [0, T_S]} and $P_S=\{P_S(t, T_S), t$ [0, T_S]} of the longterm and short-term maturity bond have a unique, continuous and square integrable solution according to Theorem 4.1, Ito's Lemma can be employed to derive the following differential equation of the relative price process $P_L^* = \{P_L^*(t, T_L) = P_L(t, T_L)/P_S(t, T_S), t \in [0, T_S]\}$,

$$\frac{dP_L^*(t,T_L)}{P_L^*(t,T_L)} = \frac{dP_L(t,T_L)}{P_L(t,T_L)} - \frac{dP_S(t,T_S)}{P_S(t,T_S)} - \frac{dP_L(t,T_L)}{P_L(t,T_L)}\frac{dP_S(t,T_S)}{P_S(t,T_S)} + \tag{4.7}$$

$$\left(\frac{dP_S(t,T_S)}{P_S(t,T_S)}\right)^2 = \mu_L^*(t)dt + \sigma_{L_1}^*(t)dW_1(t) + \sigma_{L_2}^*(t)dW_2(t)$$

with

$$\mu_L^*(t) = \mu_L(t) - \mu_S(t) + (\sigma_{S_1}^2 + \sigma_{S_2}^2)(T_S - t)^2 -$$

$$(\sigma_{L_1}\sigma_{S_1} + \sigma_{L_2}\sigma_{S_2}^2)(T_L - t)(T_S - t)$$

and

$$\sigma_{L_1}^*(t) = -\sigma_{L_1}(T_L - t) + \sigma_{S_1}(T_S - t)$$

$$\sigma_{L_2}^*(t) = -\sigma_{L_2}(T_L - t) + \sigma_{S_2}(T_S - t)$$

The following theorem asserts the existence of a unique equivalent probability measure Q under which the relative price process $P_L^* = \{P_L^*(t, T_L)$, t [0, T_S]} is a martingale.

Theorem 4.2 *The process* $\tilde{P}_L^* = \{\tilde{P}_L^*(t, T_L), t \in [0, T_S]\}$ *defined through*

$$\frac{d\tilde{P}_L^*(t, T_L)}{\tilde{P}_L^*(t, T_L)} = \sigma_{L_1}^*(t)dW_1(t) + \sigma_{L_2}^*(t)dW_2(t) \tag{4.8}$$

is a martingale on the filtered probability space $(\Omega, \mathcal{F}, F, P)$ *if the following sufficient condition holds,*

$$4T_S^2\frac{(\sigma_{L_1} - \sigma_{S_1})^2 + (\sigma_{L_2} - \sigma_{S_2})^2}{\min_{t\in[0,T_S]}\sigma_P^*(t)^2} \le \pi^2 \tag{4.9}$$

with

$$\sigma_{P^*}(t)^2 = (-\sigma_{L_1}(T_L - t) + \sigma_{S_1}(T_S - t))^2 + \qquad (4.10)$$
$$(-\sigma_{L_2}(T_L - t) + \sigma_{S_2}(T_S - t))^2$$

Furthermore, the equivalent probability measure Q *is unique and the distribution of the process* $\tilde{P}_L^* = \{\tilde{P}_L^*(t, T_L), t \in [0, T_S]\}$ *on* $(\Omega, \mathcal{F}, F, P)$ *corresponds with the distribution of the process* $P_L^* = \{P_L^*(t, T_L), t \in [0, T_S]\}$ *on* $(\Omega, \mathcal{F}, F, Q)$.

Proof See Appendix B.■

Based on this unique martingale measure[7] Q, the value at time t $[0, T_S]$ of a European call option $C^*(t, T_S, T_L, K)$ with maturity date T_S, exercise price K and written on a discount bond with maturity T_L, in terms of the value of the short-term maturity discount bond $P_S(t, T_S)$ is

$$C^*(t, T_S, T_L, K) = E_Q(Max(P_L^*(t, T_L) - K, 0) \mid \mathcal{F}_t) \qquad (4.11)$$

which yields, after applying some standard calculus,

$$C^*(t, T_S, T_L, K) = P_L^*(t, T_L)N(d_1) - KN(d_2) \qquad (4.12)$$

with

$$d_1 = \frac{1}{v}\ln\left(\frac{P^*(t, T_L)}{K}\right) + \frac{1}{2}v$$

$$d_2 = d_1 - v$$

$$v^2 = \int_t^{T_S} (\sigma_{L_1}^*(s)^2 + \sigma_{L_2}^*(s)^2)ds$$

Evaluating the last integral yields,

$$v^2 = \frac{1}{3}((T_L - t)^3 - (T_L - T_S)^3)(\sigma_{L_1}^2 + \sigma_{L_2}^2) + \frac{1}{3}(T_S - t)^3(\sigma_{S_1}^2 + \sigma_{S_2}^2) + \qquad (4.13)$$
$$\left(-T_L T_S^2 + \frac{1}{3}T_S^3 + 2tT_S T_L - t^2 T_L - t^2 T_S + \frac{2}{3}t^3\right)(\sigma_{L_1}\sigma_{S_1} + \sigma_{L_2}\sigma_{S_2})$$

The price of the European call option $C(t, T_S, T_L, K)$ at time t $[0, T_S]$ in terms of the original units of measurement is, therefore,

$$C(t, T_S, T_L, K) = P_L(t, T_L)N(d_1) - KP_S(t, T_S)N(d_2) \tag{4.14}$$

The similarity between this valuation formula and the Black-Scholes for-mula,[8] derived in Chapter 2, is striking. The option value is equal to a weighted average of the values of the long-term and short-term maturity bond. The first term in this weighted average or hedge portfolio is the current value of the long-term bond, multiplied by the hedge ratio $N(d_1)$. The second term is the total amount of borrowed money to finance the hedge position or, which is equivalent, the present value of the exercise price times the probability $N(d_2)$, under the equivalent martingale measure Q, of the option maturing in-the-money.

The volatility parameter v of the option formula reflects both the "drift-to-face-value" effect of the underlying bond as well as the stochastic behavior of the short-term maturity bond. Because the interval of possible long-term maturity bond values at maturity of the option is much more narrow compared to the range of possible stock values in the case of the Black-Scholes economy, the volatility parameter decreases at a higher rate as the maturity date nears. If the volatility of the underlying bond and the covariance between the long-term and short-term maturity bond is constant, a higher volatility of the short-term yield increases the value of the option. Because the European call option is an increasing convex function of the short-term yield, this positive relationship is easily understood.

The value of a European put option $P(t, T_S, T_L, K)$ at time t [0, T_S] with maturity date T_S, exercise price K and written on the long-term maturity discount bond with maturity T_L can be derived similarly or by making use of the put-call parity,[9] yielding

$$P(t, T_S, T_L, K) = KP_S(t, T_S)N(-d_2) - P_L(t, T_L)N(-d_1) \tag{4.15}$$

ESTIMATION OF THE MODEL

As discussed in the example on p. 31, the value of the European call option on the stock depends on the volatility of the stock. The drift term or mean return of the stock does not enter the option formula, as different investors have only to agree on the null sets of their probability measures. The actual assessment of probabilities of certain events, represented by the particular value of the drift coefficient, is the result of individual risk-return preferences and is, therefore, not important for the arbitrage-free valuation of contingent claims.

The previous section derived the value of a European call and put option on a discount bond and noted the similarity between that derivation and the Black-Scholes option pricing formula. The value of the options on a discount bond depends only on the volatilities of the underlying bond and the short-term maturity bond, and their correlation. In this section, maximum likelihood estimators of these particular parameters will be obtained, even for the case when the time intervals between the different observations are not equal. The asymptotic distribution of these estimators will, in addition, be derived. Because the standard errors of the different estimators also imply a standard error of the calculated option price, this asymptotic distribution may be of some interest for future empirical research regarding the valuation of interest rate derivative securities.[10]

Using the notation of the previous sections, suppose there is a sample of N+1 observations of the long-term and short-term maturity bond,

$$\{(B_L(t_0, T_L), B_S(t_0, T_S)), \ldots, (B_L(t_N, T_L), B_S(t_N, T_S))\}$$

with $t_0 < t_1 < \ldots < t_N$ and $t_N < T_S$. Using the relationship between these bond values and their corresponding yields, the above sample of observations can be transformed to the following sample of observations of longterm and short-term maturity yields:

$$\{(y_L(t_0), y_S(t_0)), \ldots, (y_L(t_N), y_S(t_N))\}$$

Based on the stochastic differential equation 4.2, the yields at some time t, given the information at time $s < t$ are normally distributed,[11] that is,

$$\begin{pmatrix} y_L(t) \\ y_S(t) \end{pmatrix} \mid \mathcal{F}_s \sim N(\mu_{LS}(t-s), \Omega_{LS}(t-s)) \qquad (4.16)$$

with

$$\mu_{LS}(t-s) = \begin{pmatrix} y_L(s) + \alpha_L(t-s) \\ y_S(s) + \alpha_S(t-s) \end{pmatrix}$$

and

$$\Omega_{LS}(t-s) = \begin{pmatrix} (\sigma_{L_1}^2 + \sigma_{L_2}^2)(t-s) & (\sigma_{L_1}\sigma_{S_1} + \sigma_{L_2}\sigma_{S_2})(t-s) \\ (\sigma_{L_1}\sigma_{S_1} + \sigma_{L_2}\sigma_{S_2})(t-s) & (\sigma_{S_1}^2 + \sigma_{S_2}^2)(t-s) \end{pmatrix}$$

After defining

$$\begin{aligned} \sigma_L^2 &= \sigma_{L_1}^2 + \sigma_{L_2}^2 \\ \sigma_S^2 &= \sigma_{S_1}^2 + \sigma_{S_2}^2 \\ \rho_{LS}\sigma_L\sigma_S &= \sigma_{L_1}\sigma_{S_1} + \sigma_{L_2}\sigma_{S_2} \end{aligned}$$

the likelihood function[12] can be obtained

$$\mathcal{L}(\alpha_L, \alpha_S, \sigma_L, \sigma_S, \rho_{LS} \mid \{(y_L(t_0), y_S(t_0)), \ldots, (y_L(t_N), y_S(t_N))\}) \qquad (4.17)$$

Based on this likelihood function, the Maximum Likelihood (ML) estimators of the parameters of interest are

$$\hat{\sigma}_L^2 = \frac{1}{N}\sum_{i=1}^{i=N}\frac{(y_L(t_i)-y_L(t_{i-1}))^2}{t_i-t_{i-1}} - \frac{1}{N}\frac{(y_L(t_N)-y_L(t_0))^2}{t_N-t_0} \quad (4.18)$$

$$\hat{\sigma}_S^2 = \frac{1}{N}\sum_{i=1}^{i=N}\frac{(y_S(t_i)-y_S(t_{i-1}))^2}{t_i-t_{i-1}} - \frac{1}{N}\frac{(y_S(t_N)-y_S(t_0))^2}{t_N-t_0} \quad (4.19)$$

$$\hat{\rho}_{LS} = \frac{1}{N\hat{\sigma}_L\hat{\sigma}_S}\sum_{i=1}^{i=N}\frac{(y_L(t_i)-y_L(t_{i-1}))(y_S(t_i)-y_S(t_{i-1}))}{t_i-t_{i-1}} - \frac{1}{N\hat{\sigma}_L\hat{\sigma}_S}\frac{(y_L(t_N)-y_L(t_0))(y_S(t_N)-y_S(t_0))}{t_N-t_0} \quad (4.20)$$

Under mild regularity conditions, these ML estimators are consistent, asymptotically normally distributed, and efficient in the class of all consistent and uniformly asymptotically normal (CUAN) estimators.[13]

Furthermore, the vector of estimated $(\hat{\sigma}_L^2, \hat{\sigma}_S^2, \hat{\rho}_{LS})^T$ parameter has asymptotically the following distribution:

$$\sqrt{N}\begin{pmatrix} \hat{\sigma}_L^2 - \sigma_L^2 \\ \hat{\sigma}_S^2 - \sigma_S^2 \\ \hat{\rho}_{LS} - \rho_{LS} \end{pmatrix} \overset{A}{\sim} N(0, \Omega_{ML}) \quad (4.21)$$

with

$$\Omega_{ML} = \begin{pmatrix} 2\sigma_L^4 & 2\rho_{LS}^2\sigma_L^2\sigma_S^2 & \rho_{LS}(1-\rho_{LS}^2)\sigma_L^2 \\ 2\rho_{LS}^2\sigma_L^2\sigma_S^2 & 2\sigma_S^4 & \rho_{LS}(1-\rho_{LS}^2)\sigma_S^2 \\ \rho_{LS}(1-\rho_{LS}^2)\sigma_L^2 & \rho_{LS}(1-\rho_{LS}^2)\sigma_S^2 & (1-\rho_{LS}^2)^2 \end{pmatrix}$$

and $\overset{A}{\sim}$ indicates an asymptotic relationship.[14]

In Ball and Torous (1983, Section V), the volatility parameters are estimated using logarithmic bond returns. In addition, it is assumed that the mean logarithmic bond return at some time t_i, i=1, …, N is equal to the expected return at time t_0. Although the proposed estimators are unbiased if the time interval between different observations approaches zero, the estimation procedure derived above is preferable, as no additional assumptions about the observed sample of bond prices are necessary.

As already mentioned in the introduction to this chapter, the derivation of the precise estimation of this model is supposed to enhance the understanding of the current valuation model. The partial equilibrium between the long-term and short-term bond which is assumed to exist and excludes the possibility of arbitrage opportunities, allows for the valuation of options on discount bonds. This approach is similar to the model of

Black and Scholes and it is interesting to compare the results obtained thus far with the general equilibrium results of the next chapter. A practical implementation of this model, however, is severely limited by the fact that options only on coupon paying bonds are traded. Because a direct stochastic description of these bonds incorporating the additional price characteristics is not possible yet,[15] and the corresponding derivative securities cannot be valued, no further attention will be given to an empirical investigation of these approaches.

APPENDIX A

Suppose $x(0) \in \mathbb{R}^N$, $\mu(x, t)$ is a $\mathbb{R}^N$-valued function that is measurable with respect to all its arguments and $\sigma(x, t)$ is a $\mathbb{R}^{N \times M}$-valued function that is also measurable with respect to all its arguments. According to Gihman and Skorohod (1972, p. 40), the stochastic differential equation given by

$$dx(t) = \mu(x, t)dt + \sigma(x, t)dW(t) \tag{A.1}$$

with $W=\{W_1(t),\ldots, W_M(t), t \quad [0, T_S]\}$ *an* M-dimensional Standard Brownian Motion initialized at zero, has a unique, continuous and square integrable solution $x = \{x(t), t \quad [0, T]\}$, if there exists a constant $C \in \mathbb{R}$ such that

$$\|\mu(x, t) - \mu(y, t)\| + \|\sigma(x, t) - \sigma(y, t)\| \le C\|x - y\| \tag{A.2}$$

and

$$\|\mu(x, t)\|^2 + \|\sigma(x, t)\|^2 \le C(1 + \|x\|^2) \tag{A.3}$$

holds true for all $t \quad [0, T]$ and $x, y \in \mathbb{R}^N$ with the Euclidean norm $\|A\| = (tr(AAT))^{½}$ for $A \in \mathbb{R}^{N \times M}$.

To prove that the stochastic differential equation given by equation (4.3) has a solution with the above-mentioned properties, it is sufficient to show that the following augmented system of differential equations fulfils (A.2) and (A.3),

$$\begin{pmatrix} dW_1(t) \\ dW_2(t) \\ dW_3(t) \\ dW_4(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \mu_{W_3}(t) \\ \mu_{W_4}(t) \end{pmatrix} dt + \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ & -\Omega(t) & \end{pmatrix} \begin{pmatrix} dW_1(t) \\ dW_2(t) \end{pmatrix} \tag{A.4}$$

with $W_3(t) = lnP_L(t, T_L)$, $W_4(t) = lnP_S(t, T_S)$. Using Arnold (1974, p. 142) equation (4.2) can be written as

$$\begin{pmatrix} y_L(t) \\ y_S(t) \end{pmatrix} = \begin{pmatrix} \alpha_L \\ \alpha_S \end{pmatrix} t + \begin{pmatrix} \sigma_{L_1} & \sigma_{L2} \\ \sigma_{S_1} & \sigma_{S_2} \end{pmatrix} \begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix} \tag{A.5}$$

which allows the specification of the drift functions $\mu w_3(t)$ and $\mu w_4(t)$ in terms of $W_1(t)$ and $W_2(t)$, *as follows:*

$$\begin{pmatrix} \mu_{W_3}(t) \\ \mu_{W_4}(t) \end{pmatrix} = \begin{pmatrix} y_L(0) + \alpha_L t - \alpha_L(T_L - t) \\ y_S(0) + \alpha_S t - \alpha_S(T_S - t) \end{pmatrix} + \begin{pmatrix} \sigma_{L_1} & \sigma_{L_2} \\ \sigma_{S_1} & \sigma_{S_2} \end{pmatrix} \begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix} \tag{A.6}$$

Define now

$$x(t) = \begin{pmatrix} W_1(t) \\ W_2(t) \\ W_3(t) \\ W_4(t) \end{pmatrix}; \ \mu(x, t) = \begin{pmatrix} 0 \\ 0 \\ \mu_{W_3}(t) \\ \mu_{W_4}(t) \end{pmatrix}; \ \sigma(x, t) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ & -\Omega(t) \end{pmatrix}$$

Because $\mu(x, t)$ is linear in $x(t)$ and $\sigma(x, t)$ is independent of $x(t)$, a constant $C_1 \in \mathbb{R}$ can easily be found such that equation (A.2) is fulfilled for all t [0, Ts] and all $x, y \in \mathbb{R}^4$.

To prove condition (A.3), now define

$$a(t) = \begin{pmatrix} y_L(0) + \alpha_L t - (T_L - t)\alpha_L \\ y_S(0) + \alpha_S t - (T_S - t)\alpha_S \end{pmatrix}$$

and

$$B(t) = \begin{pmatrix} 0 & 0 & \sigma_{L_1} & \sigma_{L_2} \\ 0 & 0 & \sigma_{S_1} & \sigma_{S_2} \end{pmatrix}$$

after which this can be written

$$\|\mu(x, t)\|^2 + \|\sigma(x, t)\|^2 = \|a(t) + B(t)x\|^2 + \|\Omega(t)\|^2 + 4 \tag{A.7}$$

for all t [0, T_S] and all $x \in \mathbb{R}^4$. Taking appropriate real-valued constants C_2, C_3 and C_4 yields

$$\|a(t) + B(t)x\|^2 + \|\Omega(t)\|^2 + 4 \leq C_2 + C_3\|x\| + C_4\|x\|^2 \tag{A.8}$$

For the case $\|x\| > 1$,

$$C_2 + C_3\|x\| + C_4\|x\|^2 \le C_5(1 + \|x\|^2) \tag{A.9}$$

with $C_5=Max(C_2, C_3+C_4)$.

For the case $\|x\| < 1$, simply take $C_5=C_2+C_3+C_4$ to get

$$C_2 + C_3\|x\| + C_4\|x\|^2 \le C_5 \tag{A.10}$$

By taking C=$Max(C_1, C_5)$, finally, the desired result has been obtained.

APPENDIX B

The first step of the proof[16] consists of rewriting the stochastic differentialequation of the relative price as $P_L^* = \{P_L^*(t, T_L), t \in [0, T_S]\}$

$$\frac{dP_L^*(t, T_L)}{P_L^*(t, T_L)} = \mu_{P^*}(Z_1(t), t)dt + \sigma_{P^*}(t)dZ_2(t) \tag{B.l}$$

In this equation, the drift and volatility functions are equal to

$$\mu_{P^*}(Z_1(t), t) = y_L(0) + \alpha_L t - \alpha_L(T_L - t) - y_S(0) - \alpha_S t +$$

$$\alpha_S(T_S - t) + \sqrt{(\sigma_{L_1} - \sigma_{S_1})^2 + (\sigma_{L_2} - \sigma_{S_2})^2}Z_1(t) +$$

$$\frac{1}{2}(\sigma_{L_1}^2 + \sigma_{L_2}^2)(T_L - t)^2 + \frac{1}{2}(\sigma_{S_1}^2 + \sigma_{S_2}^2)(T_S - t)^2 -$$

$$(\sigma_{L_1}\sigma_{S_1} + \sigma_{L_2}\sigma_{S_2})(T_L - t)(T_S - t)$$

and

$$\sigma_{P^*}(t)^2 = (-\sigma_{L_1}(T_L - t) + \sigma_{S_1}(T_S - t))^2 +$$

$$(-\sigma_{L_2}(T_L - t) + \sigma_{S_2}(T_S - t))^2$$

The Brownian Motions $(Z_1, Z_2)=\{(Z_1(t), Z_2(t)), t \quad [0, T_S]\}$ are defined such that

$$\sqrt{(\sigma_{L_1} - \sigma_{S_1})^2 + (\sigma_{L_2} - \sigma_{S_2})^2}Z_1(t) = (\sigma_{L_1} - \sigma_{S_1})W_1(t) + \tag{B.2}$$

$$(\sigma_{L_2} - \sigma_{S_2})W_2(t)$$

and

$$\sigma_{P^*}(t)dZ_2(t) = (-\sigma_{L_1}(T_L - t) + \sigma_{S_1}(T_S - t))dW_1(t) + (-\sigma_{L_2}(T_L - t) + \sigma_{S_2}(T_S - t))dW_2(t) \quad \text{(B.3)}$$

In order to apply Girsanov's Theorem to obtain a unique equivalent probability *measure* Q such that the relative price process $P_L^* = \{P_L^*(t, T_L)$, t [0, T_S]} is a martingale with respect to this measure, the Radon-Nikodym derivative which denotes the corresponding change of probability measure given by

$$\rho(T_S) = exp\left(\int_0^{T_S} \alpha(Z_1(t), t)dZ_2(t) - \frac{1}{2}\int_0^{T_S} \alpha(Z_1(t), t)^2 dt\right) \quad \text{(B.4)}$$

with

$$\alpha(Z_1(t), t) = -\frac{\mu_{P^*}(Z_1(t), t)}{\sigma_{P^*}(t)} \quad \text{(B.5)}$$

must be well-defined. A sufficient condition to guarantee the existence of such a unique equivalent probability measure, as stated in Müller (1985, p. 91) and Cheng (1989, pp. 188–9), is the following inequality

$$E_P\left[exp\left(\frac{1}{2}\int_0^{T_S} \alpha(Z_1(t), t)^2 dt\right)\right] < \infty \quad \text{(B.6)}$$

To investigate the conditions under which this inequality holds, first define

$$f_1 = \sqrt{(\sigma_{L_1} - \sigma_{S_1})^2 + (\sigma_{L_2} - \sigma_{S_2})^2} \quad \text{(B.7)}$$

and

$$f_2(t) = \mu_{P^*}(Z_1(t), t) - f_1 Z_1(t) \quad \text{(B.8)}$$

to obtain

$$\alpha(Z_1(t), t) = -\frac{f_1 Z_1(t) + f_2(t)}{\sigma_{P^*}(t)} \quad \text{(B.9)}$$

Obviously, there is less than perfect correlation between the two bond price processes $P_L=\{P_L(t, T_L), t$ [0, T_S]} and $P_s=\{P_S(t, T_S), t$ [0, T_S]}.

Therefore

$$\sigma_{L_1}\sigma_{S_2} \neq \sigma_{L_2}\sigma_{S_1} \quad \text{(B.10)}$$

Moreover, there is

$$\exists \epsilon_\rho > 0 \quad \text{such that} \quad \sigma_{P\bullet}^2(t)^2 \geq \epsilon_\rho$$

for all t [0, T_S]. This inequality can be used to write

$$\int_0^{T_S} \alpha(Z_1(t), t)^2 dt \leq \frac{f_1^2}{\epsilon_\rho} \int_0^{T_S} Z_1^2(t)dt + 2\frac{f_1}{\epsilon_\rho} \int_0^{T_S} f_2(t)Z_1(t)dt + \tag{B.11}$$

$$\frac{1}{\epsilon_\rho} \int_0^{T_S} f_2^2(t)dt$$

Because the function $f_2(t)$ is uniformly bounded and continuous on [0, T_S]:

$$\exists C \in \mathbb{R} \quad \text{such that} \quad \int_0^{T_S} f_2^2(t)dt \leq C$$

Denoting $f_2(t)=dF_2(t)$, yields

$$\int_0^{T_S} f_2(t)Z_1(t)dt = \int_0^{T_S} Z_1(t)dF_2(t) = \tag{B.12}$$

$$Z_1(t)F_2(t)\,\big|_0^{T_S} - \int_0^{T_S} F_2(t)dZ_1(t) =$$

$$Z_1(T_S)F_2(T_S) - Z_1(0)F_2(0) - \int_0^{T_S} F_2(t)dZ_1(t)$$

Denoting this expression by $\tilde{X}$ yields

$$\tilde{X} \sim N\left(Z_1(T_S)F_2(T_S) - Z_1(0)F_2(0), \int_0^{T_S} F_2^2(t)dt\right) \tag{B.13}$$

Therefore

$$\int_0^{T_S} \alpha(Z_1(t), t)^2 dt \leq \frac{1}{\epsilon_\rho} C + 2\frac{f_1}{\epsilon_\rho}\tilde{X} + \frac{f_1^2}{\epsilon_\rho} \int_0^{T_S} Z_1^2(t)dt \tag{B.14}$$

so

$$E_P\left[exp\left(\frac{1}{2}\int_0^{T_S}\alpha(Z_1(t),t)^2dt\right)\right] \leq \tag{B.15}$$

$$exp\left(\frac{1}{2\epsilon_\rho}C\right)E_P\left[exp\left(\frac{f_1}{\epsilon_\rho}\tilde{X}\right)exp\left(\frac{f_1^2}{2\epsilon_\rho}\int_0^{T_S}Z_1^2(t)dt\right)\right]$$

Define now

$$\tilde{Y}_1 = exp\left(\frac{f_1}{\epsilon_\rho}\tilde{X}\right) \tag{B.16}$$

and

$$\tilde{Y}_2 = exp\left(\frac{f_1^2}{2\epsilon_\rho}\int_0^{T_S}Z_1^2(t)dt\right) \tag{B.17}$$

It must be shown that

$$E_P[\tilde{Y}_1\tilde{Y}_2] < \infty \tag{B.18}$$

Because

$$E_P[\tilde{Y}_1\tilde{Y}_2] \leq \sqrt{(E_P[\tilde{Y}_1^2]-E_P[\tilde{Y}_1]^2)(E_P[\tilde{Y}_2^2]-E_P[\tilde{Y}_2]^2)}+ \tag{B.19}$$

$$E_P[\tilde{Y}_1]E_P[\tilde{Y}_2]$$

and $(\tilde{Y}_1)$ is normally distributed with finite mean and variance, it remains to be shown that $E_P[\tilde{Y}_2^2]$ and $E_P[\tilde{Y}_2]$ are finite. Consider therefore, the following expectation:

$$E_P[\tilde{Y}_2^\beta] = E_P\left[exp\left(\gamma\int_0^{T_S}Z_1^2(t)dt\right)\right] \tag{B.20}$$

with

$$\gamma = \frac{2\beta\epsilon_\rho}{f_1^2}$$

In the remaining part of this proof, this integral will be written as the limit of the following Riemann-summation, that is,

$$\int_0^{T_S} Z_1^2(t)dt = \sum_{i=1}^{i=N} Z_1^2(t_i)(t_i - t_{i-1}) = \frac{T_S}{N}\sum_{i=1}^{i=N} Z_1^2(t_i) \quad \text{(B.21)}$$

with an equidistant partition and $t_0=0$ and $t_N=T_S$. The vector of stochastic variables $Z_1=(Z_1(t_0),\ldots, Z_1(t_N))^T$ is normally distributed with zero mean and covariance-matrix

$$\boldsymbol{\Omega}_N = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2 & \cdots & 2 \\ 1 & 2 & 3 & \cdots & 3 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & 2 & 3 & \cdots & N \end{pmatrix}\left(\frac{T_S}{N}\right) \quad \text{(B.22)}$$

The inverted matrix $\boldsymbol{\Omega}_N^{-1}$ is equal to

$$\boldsymbol{\Omega}_N^{-1} = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}\left(\frac{N}{T_S}\right) \quad \text{(B.23)}$$

which can be used to solve the expectation

$$E_P\left[exp\left(\gamma\int_0^{T_S} Z_1^2(t)dt\right)\right] = \lim_{N\to\infty} E_P\left[exp\left(\gamma\frac{T_S}{N}\sum_{i=1}^{i=N} Z_1^2(t_i)\right)\right] = \quad \text{(B.24)}$$

$$\lim_{N\to\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(2\pi)^{-\frac{N}{2}}\,|\,\boldsymbol{\Omega}_N\,|^{-\frac{1}{2}}\,exp\left(\gamma\frac{T_S}{N}\boldsymbol{Z}_1^T\boldsymbol{Z}_1 - \frac{1}{2}\boldsymbol{Z}_1^T\boldsymbol{\Omega}_N^{-1}\boldsymbol{Z}_1\right)\times$$

$$dZ_1(t_1)\cdots dZ_1(t_N) = \lim_{N\to\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(2\pi)^{-\frac{N}{2}}\,|\,\boldsymbol{\Omega}_N\,|^{-\frac{1}{2}}\times$$

$$exp\left(-\frac{1}{2}\boldsymbol{Z}_1^T\left(\boldsymbol{\Omega}_N^{-1} - 2\gamma\frac{T_S}{N}\boldsymbol{I}_N\right)\boldsymbol{Z}_1\right)dZ_1(t_1)\cdots dZ_1(t_N)$$

It is obvious that | . | denotes the determinant of a matrix. Suppose $\boldsymbol{V}_N^{-1} = \boldsymbol{\Omega}_N^{-1} - 2\gamma\frac{T_S}{N}\boldsymbol{I}_N$. If V_N is positive definite for all N, this results in,

$$\lim_{N\to\infty}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}(2\pi)^{-\frac{N}{2}}\,|V_N|^{-\frac{1}{2}}\,exp\left(-\frac{1}{2}Z_1^T V_N Z_1\right)\times \tag{B.25}$$

$$dZ_1(t_1)\cdots dZ_1(t_N)\times |V_N|^{\frac{1}{2}}|\Omega_N|^{-\frac{1}{2}} = \lim_{N\to\infty}|V_N|^{\frac{1}{2}}|\Omega_N|^{-\frac{1}{2}} =$$

$$\lim_{N\to\infty}|V_N^{-1}|^{-\frac{1}{2}}|\Omega_N|^{-\frac{1}{2}} = \lim_{N\to\infty}|V_N^{-1}|^{-\frac{1}{2}}\left(\frac{N}{T_S}\right)^{\frac{N}{2}} =$$

$$\lim_{N\to\infty}|V_N^{-1}|^{-\frac{1}{2}}\left(\frac{T_S}{N}\right)^{-\frac{N}{2}} = \lim_{N\to\infty}\left|V_N^{-1}\left(\frac{T_S}{N}\right)\right|^{-\frac{1}{2}}$$

Define

$$D_N = \left|V_N^{-1}\left(\frac{T_S}{N}\right)\right| = \left|\left(\Omega_N^{-1} - 2\gamma\frac{T_S}{N}I_N\right)\left(\frac{T_S}{N}\right)\right| \tag{B.26}$$

then

$$D_N = \begin{vmatrix} 2-2\gamma\left(\frac{T_S}{N}\right)^2 & -1 & 0 & \cdots & 0 \\ -1 & 2-2\gamma\left(\frac{T_S}{N}\right)^2 & -1 & \cdots & 0 \\ 0 & -1 & 2-2\gamma\left(\frac{T_S}{N}\right)^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & -1 & 1-2\gamma\left(\frac{T_S}{N}\right)^2 \end{vmatrix} \tag{B.27}$$

This specific structure allows for the derivation of the following recursive relationship

$$D_k = \left(2-2\gamma\left(\frac{T_S}{N}\right)^2\right)D_{k-1} - D_{k-2} \qquad k = 2, 3, \ldots, N \tag{B.28}$$

with

$$D_0 = 1 \tag{B.29}$$

and

$$D_1 = 1 - 2\gamma\left(\frac{T_S}{N}\right)^2 \tag{B.30}$$

In order to solve the difference equation, the roots of the corresponding second order polynomial must be determined:

$$\lambda^2 - 2\left(1 - 2\gamma\left(\frac{T_S}{N}\right)^2\right)\lambda + 1 = 0 \tag{B.31}$$

yielding the following two solutions

$$\lambda_1 = 1 - \epsilon_N + i\sqrt{1 - (1 - \epsilon_N)^2} \tag{B.32}$$

$$\lambda_2 = 1 - \epsilon_N - i\sqrt{1 - (1 - \epsilon_N)^2} \tag{B.33}$$

where

$$\epsilon_N = \gamma\left(\frac{T_S}{N}\right)^2$$

This can also be written as

$$\lambda_1 = cos(b) + sin(b)i \tag{B.34}$$

$$\lambda_2 = cos(b) - sin(b)i \tag{B.35}$$

The general solution to the difference equation is now

$$D_k = a_N cos(bk) + b_N sin(bk) \tag{B.36}$$

with initial conditions (B.29) and (B.30). These conditions give

$$a_N = 1 \tag{B.37}$$

and

$$b_N = -\frac{\epsilon_N}{\sqrt{1 - (1 - \epsilon_N)^2}} \tag{B.38}$$

Because the expectation is equal to

$$E_P\left[exp\left(\gamma \int_0^{T_S} Z_1^2(t)dt\right)\right] = \lim_{N\to\infty} D_N^{-\frac{1}{2}} \tag{B.39}$$

we first determine

$$\lim_{N\to\infty} D_N = \lim_{N\to\infty} cos(bN) - \frac{\epsilon_N}{\sqrt{1-(1-\epsilon_N)^2}} sin(bN) \tag{B.40}$$

It is known that parameter *b* fulfils

$$cos(b) = 1 - \epsilon_N \tag{B.41}$$

and

$$sin(b) = \sqrt{1-(1-\epsilon_N)^2} \tag{B.42}$$

yielding

$$b = arctan\left(\sqrt{\frac{1}{(1-\epsilon_N)^2} - 1}\right) =: arctan(x_N) \tag{B.43}$$

with $0 < x_N < 1$. This limit can be solved by using an upper and lower boundary of *arctan(*x_N*)*, that is,

$$N\left(x_N - \frac{1}{3}x_N^3\right) \le N\,arctan(x_N) \le N x_N \tag{B.44}$$

which gives

$$\sqrt{2\gamma T_S^2} \le \lim_{N\to\infty} N\,arctan(x_N) \le \sqrt{2\gamma T_S^2} \tag{B.45}$$

The above transformation of multivariate normal distributions is allowed only if the matrix V_N is positive definite. Because D_k, $k = 1, 2, \ldots$, is monotonically decreasing and

$$\lim_{N\to\infty} D_N > 0 \quad \text{if} \quad \sqrt{2\gamma T_S^2} \le \frac{\pi}{2} \tag{B.46}$$

the matrix V_N is positive definite. Therefore

$$\lim_{N\to\infty} D_N^{-\frac{1}{2}} = \frac{1}{\sqrt{cos\left(\sqrt{2\gamma T_S^2}\right)}} \quad \text{If} \quad \sqrt{2\gamma T_S^2} \le \frac{\pi}{2} \tag{B.47}$$

Because it has been shown that a sufficient condition is fulfilled for the existence of a unique equivalent martingale measure if

$$\sqrt{2\gamma T_S^2} \leq \frac{\pi}{2} \qquad \text{(B.48)}$$

or, equivalently,

$$4T_S^2 \frac{(\sigma_{L_1} - \sigma_{S_1})^2 + (\sigma_{L_2} - \sigma_{S_2})^2}{\epsilon_D} \leq \pi^2 \qquad \text{(B.49)}$$

it is hard to investigate this existence if (B.49) does not hold. As such, this situation will not be further analyzed here.

NOTES

1 Based on these specific bond characteristics, rational boundaries, which option prices have to obey, can be derived. Throughout this thesis, different models are frequently illustrated by means of these boundaries. For an overview, see Rady and Sandmann (1992).

2 In Schöbel (1986), this model is extended by deriving option values subject to the constraint of long-term bonds having a lower value than short-term bonds, that is, positive forward rates. The resulting option formula contains an additional term, called the "anti-option," which can be interpreted as a discount to the regular option value, as negative forward rates are excluded. However, the original discount bond price processes still allow for negative forward rates. Only the option value has been obtained subject to this constraint. Because of this inconsistency, no further attention will be paid to this extended valuation approach.

3 See, for example, Gihman and Skorohod (1972, p. 269–70).

4 See again Gihman and Skorohod (1972, p. 40).

5 This equation is obtained simply by rewriting equation (7) of Ball and Torous (1983, p. 526).

6 As a result of this small effective probability of negative interest rates, the value of an option on a discount bond with a strike price equal to the face value of the underlying bond is insignificant for reasonable values of the necessary parameters. For this reason, no attention will be paid to the model of Bühler and Käsler (1989). Although their model excludes negative interest rates, the stochastic differential equations of the bond prices do not permit any empirical analysis.

7 The sufficient condition regarding the existence of the unique equivalent martingale measure is not stringent. Suppose the following reasonable values for the different parameters: $\sigma_{L_1} = 0.02,\ \sigma_{L_2} = 0.02,\ \sigma_{S_1} = 0.03,\ \sigma_{S_2} = 0.03,$ T_L=10 and T_S=5. It is easy to verify that the sufficient condition holds.

8 Although this thesis is concerned with the theoretical and empirical valuation of interest rate derivative securities, it is interesting to apply the analysis of this chapter to the valuation of options on stocks. Suppose S(t denotes the value of the stock at time t [0, T_S]. Assuming the following relationship between the long term yield and the stock,

$$S(t) = e^{y_L(t)}$$

it is easily verified that Theorems 4.1 and 4.2 apply to the resulting price processes as well. The value of a European call option *C(t, T_S, S(t), K)* on the stock can be derived similarly and is equal to

$$C(t, T_S, S(t), K) = S(t)N(d_1) - KP_S(t, T_S)N(d_2)$$

with

$$d_1 = \frac{1}{v} \ln\left(\frac{S(t)}{KP_S(t, T_S)}\right) + \frac{1}{2} v$$

$$d_2 = d_1 - v$$

$$v^2 = (T_S - t)(\sigma_{L_1}^2 + \sigma_{L_2}^2) + \frac{1}{3}(T_S - t)^3(\sigma_{S_1}^2 + \sigma_{S_2}^2) +$$

$$(T_S - t)^2(\sigma_{L_1}\sigma_{S_1} + \sigma_{L_2}\sigma_{S_2})$$

If the volatility of the short-term yield is equal to zero, we again have the familiar Black and Scholes formula.

9 The put-call parity gives the relation between a European call and put option with the same exercise price and the same maturity date, that is,

$$P_L(t, T_L) - KP_S(t, T_S) = C(t, T_S, T_L, K) - P(t, T_S, T_L, K)$$

The value of a long call and short put is equal to the underlying value less the present value of the exercise price, a result that is obtained simply by comparing the payout of both positions at maturity of the option and taking the expectation of both positions under the probability measure Q. Using the relationship

$$N(-x) = 1 - N(x)$$

and the call option valuation formula, the put option valuation formula is directly obtained.

10 In Lo (1986), for example, extensive simulations are carried out to determine the minimum number of observations necessary to be able to apply the asymp-totic theory in the case of the Black and Scholes formula.

11 See Arnold (1974, ch. 8).

12 See, for example, Judge *et al.* (1982, Section 7.1).

13 See Kendall and Stuart (1979).

14 See again Kendall and Stuart (1979, p. 59).

15 See, for example, Rady and Sandmann (1992, p. 1).

16 We would like to thank Marc Yor for helpful comments by means of private communication regarding this proof.

5
MODELLING THE TERM STRUCTURE OF INTEREST RATES

In the indirect one-factor approach, the prices of all interest rate dependent securities are assumed to be functions of the instantaneous spot rate. The explicit description of the stochastic process of this short rate, together with the well-known conditions that prohibit arbitrage opportunities, then provide a general valuation relationship that all securities have to obey. As part of this equilibrium, the term structure of interest rates at some valuation date can be obtained. Because the stochastic process of the short rate also implies a stochastic movement of the term structure over time, a corresponding term structure of interest rate volatilities can be derived.

In Chapter 3, the various interest rates models are generally classified according to these two term structures. In this chapter, these models will be discussed in detail in the order of this classification. The first section, therefore, starts with the endogenous term structure of interest rates models. In the second section, the exogenous term structure models are explained. To maintain the classification, this section is divided into two parts. The first subsection deals with the endogenous term structure of volatilities models, while the second subsection focuses on the exogenous volatility structure models.

ENDOGENOUS TERM STRUCTURE OF INTEREST RATES MODELS

The endogenous term structures of interest rates models are principally characterized by the stochastic process of the instantaneous shortterm interest rate. The parameters of the process are assumed to be time-invariant and, therefore, no explicit reference is being made at this moment to an observed term structure of interest rates at some valuation date.

This section discusses the well-known interest rates models of Merton (1973), Vasicek (1977) and Cox, Ingersoll and Ross (1985). As will be seen in the next section, these models have a nice counterpart within the class of exogenous term structure models, which explains the detailed treatment. In addition, a great deal of attention in the empirical part of this thesis will be paid to the Vasicek and Cox *et al.* models; this section, therefore, paves the way by deriving some relations that will be estimated later on.

The description of each of these models will start with the stochastic process of the short rate. After the properties of this short rate have been discussed, both the implied term structure of interest rates and the corresponding term structure of interest rate volatilities will be derived. The value of a European call option on a discount bond will also be obtained. Although these models allow for the valuation of any interest rate dependent security according to the valuation techniques of Chapter 2, the focus here is on this particular interest rate contingent claim for two reasons.

The problem of the valuation of a call option on a discount bond has attracted a great deal of attention within the field of valuing interest rate derivative securities, as it has a clear analogue to the familiar valuation problem of options on stocks. To be in line with these well-known approaches, it is natural to concentrate on this claim.

In addition, these models have the common property that closed-form solutions can be obtained in case of call options on discount bonds. As such, these interest rates models can be compared analytically and the differences between stock and bond options given emphasis.

After this detailed explanation, some other interest rates models will be discussed. As the basic drawback of these models is their analytic complexity, which limits an economic interpretation and empirical investigation, a brief explanation of some of the basic characteristics of these models will complete the overview.

The interest rates models to be discussed in this section are one-factor interest rates models, which can easily be extended to multi-factor models. Although the complexity of the stochastic dynamics definitely increases, the main characteristics by which one compares and illustrates the different models remain the same and because of this, limited attention will be given to these multi-factor analogues.

The Merton Model

The continuous time uncertainty will be specified by the filtered probability space $(\Omega, \mathcal{F}, F, P)$, satisfying the usual conditions. As before, investors have only to agree on the null sets of the probability measure instead of an actual assessment of probabilities of certain events, which means that the probability measure *P* can be replaced by any other equivalent measure *P**.

The stochastic differential equation of the instantaneous short-term rate of interest is given in Merton (1973, p. 163) by

$$dr(t) = \theta dt + \sigma dW(t) \tag{5.1}$$

In this equation, θ and σ are real-valued constants and $W=\{W(t),\ t \geq 0\}$ is a Standard Brownian Motion. According to Arnold (1974), the short rate at time *t*, given the information set at time s, $s \leq t$, is normally distributed, that is,

$$r(t) \mid \mathcal{F}_s \sim N(r(s) + \theta(t-s), \sigma^2(t-s)) \tag{5.2}$$

It is obvious that this specification of the short rate is only expository. As already noted by Merton (1973), the assumed normality assigns positive probabilities to negative rates. In addition, if $\theta \neq 0$, the conditional mean of the short rate at time *s* increases without bound as *t* increases.[1]

In order to derive the value of an interest rate derivative security, a unique equivalent probability measure must first be identified. A stochastic process or Radon-Nikodym derivative $p=\{p(\lambda, t),\ t \geq 0\}$, is, therefore, defined by

$$\rho(\lambda, t) = exp\left(\int_0^t \lambda dW(s) - \frac{1}{2}\int_0^t \lambda^2 ds\right) \quad (5.3)$$

with λ a fixed real-valued constant. In the remainder of this section, it will be seen that A. can be interpreted as the market price of risk. By Girsanov's Theorem, the stochastic process $W = \{W(t), t \geq 0\}$, defined by

$$\tilde{W}(t) = W(t) - \int_0^t \lambda ds \quad (5.4)$$

is a Standard Brownian Motion on the probability space $(\Omega, \mathcal{F}, F, Q)$, where the unique equivalent probability measure Q is given by

$$dQ = \rho(\lambda, t)dP \quad (5.5)$$

According to Chapter 2, the value of a discount bond *P(r(t), t, τ)* at time *t* with remaining time to maturity *τ* and unit face value, equals its discounted expected value. The expectation has to be taken with respect to a unique equivalent probability measure such that the values of trading strategies in terms of the money market account are martingales. In case of the endogenous term structure models, this can be established by calculating the expectation under the unique measure Q, yielding

$$P(r(t), t, \tau) = E_Q\left(e^{-\int_t^{t+\tau} r(s)ds} \mid \mathcal{F}_t\right) = \quad (5.6)$$

$$exp\left(-r(t)\tau - (\theta + \lambda\sigma)\frac{\tau^2}{2} + \sigma^2\frac{\tau^3}{6}\right)$$

It is easy to verify that under this measure Q, investors are risk-neutral. The instantaneous expected return and volatility of this return on a zerocoupon bond under this measure are $r(t)$ and $\sigma\tau$, respectively. Under the original measure *P*, however, the instantaneous expected return and volatility are equal to $r(t)+\lambda\sigma\tau$ and $\sigma\tau$, from which λ can be interpreted as the market price of risk. If investors are risk averse and the market price of risk is positive, accordingly, the relationship between the instantaneous expected return and volatility of bonds is linear with positive slope coefficient λ.

All investors have to agree on Q and on the resulting values of discount bonds in order to exclude riskless arbitrage opportunities. Looking at the relationships between bond values and the interest rate process, this translates into an agreement between investors on the particular values of the parameters $\theta+\lambda\sigma$ and σ. It is not necessary, therefore, that the instantaneous returns on discount bonds or the market prices of risks are the same among investors.

Obtained from the values of these bonds, the endogenous term structure of interest rates $R(t, \tau)$ at time t is,

$$R(t,\tau) = -\frac{\ln P(r(t), t, \tau)}{\tau} = r(t) + (\theta + \lambda\sigma)\frac{\tau}{2} - \sigma^2\frac{\tau^2}{6} \tag{5.7}$$

This relationship allows for some interesting observations. Because the term structure of interest rates equals the sum of the instantaneous short rate and a quadratic function in τ, changes in the value of this short rate cause parallel shifts in the term structure. In addition, yields are a concave function of the volatility of the short rate. Figure 5.1 shows the term structure of interest rates for different values of this volatility. The instantaneous short rate is 0.07 and the risk-neutral drift term of the corresponding process, 0.02. The different values of the volatility are 0.01, 0.04 and 0.07, respectively.

This figure reveals that yields are negatively related to the value of the volatility. As volatility increases, the curvature of the term structure also increases and, at last, although this is not clear from Figure 5.1, the value of the infinite maturity yield $R(t, \infty)$, equals

$$R(t,\infty) = \lim_{\tau\to\infty} R(t,\tau) = -\infty \tag{5.8}$$

These relationships have a simple explanation. Due to the simple normal distribution of instantaneous spot rates, a higher volatility increases the probability of both higher and lower spot rates. The relationship, however, between bond values and spot rates is convex, implying a greater impact of lower rates on bond values. Because the volatility of spot rates also increases as a function of maturity, the above relationships are easily understood.

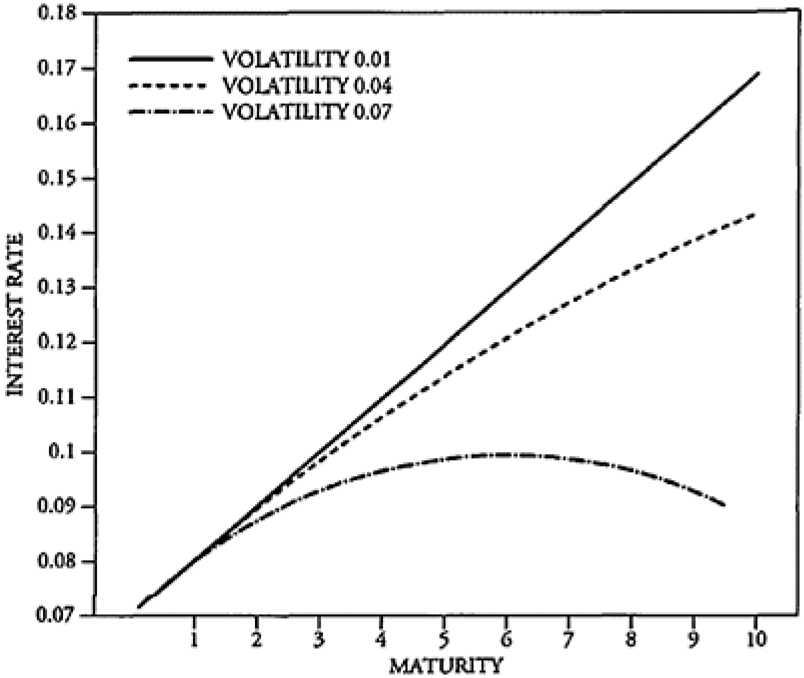

Figure 5.1 The term structure of interest rates

This figure shows the term structure of interest rates according to the model of Merton for different values of the volatilities. The instantaneous short-term rate of interest is 0.07 and the risk-neutral drift term of the corresponding process 0.02. The values of the volatility are 0.01, 0.04 and 0.07, respectively.

The yield-to-maturity $R(t, \tau)$ at time t, given the information set at time s, $s \leq t$, is normally distributed,

$$R(t, \tau) \mid \mathcal{F}_s \sim N(R(s, \tau) + \theta(t - s), \sigma^2(t - s)) \quad (5.9)$$

The volatility of the yield-to-maturity is independent of the time-tomaturity. In the classification scheme, this means that the endogenous term structure of interest rate volatilities is a flat function.

The value of a European call option on a discount bond can also be derived by taking the discounted expectation with respect to Q of the payout of the claim at maturity. Denote the value at time t of a call option with maturity τ_1 exercise price K and underlying discount bond with maturity τ_2, by $C(r(t), t, K, \tau_1, \tau_2)$; this means that

$$C(r(t), t, K, \tau_1, \tau_2) = E_Q\left(e^{-\int_t^{t+\tau_1} r(s)ds} Max(P(r(\tau_1), \tau_1, \tau_2 - \tau_1) - K, 0)\right) \quad (5.10)$$

which is equal to

$$C(r(t), t, K, \tau_1, \tau_2) = P(r(t), t, \tau_2)N(d_1) - KP(r(t), t, \tau_1)N(d_2) \quad (5.11)$$

with

$$d_1 = \frac{1}{v}\ln\left(\frac{P(r(t), t, \tau_2)}{KP(r(t), t, \tau_1)}\right) + \frac{1}{2}v$$

$$d_2 = d_1 - v$$

$$v^2 = \sigma^2(\tau_2 - \tau_1)^2(\tau_1 - t)$$

Figure 5.2 shows the value of this European call option for different maturities. The instantaneous short rate is again 0.07, while the risk-neutral drift term and volatility are 0.02 and 0.07, respectively. The maturity of the underlying bond is ten years and the exercise price is equal to the forward price of the underlying bond in order to be able to concentrate on

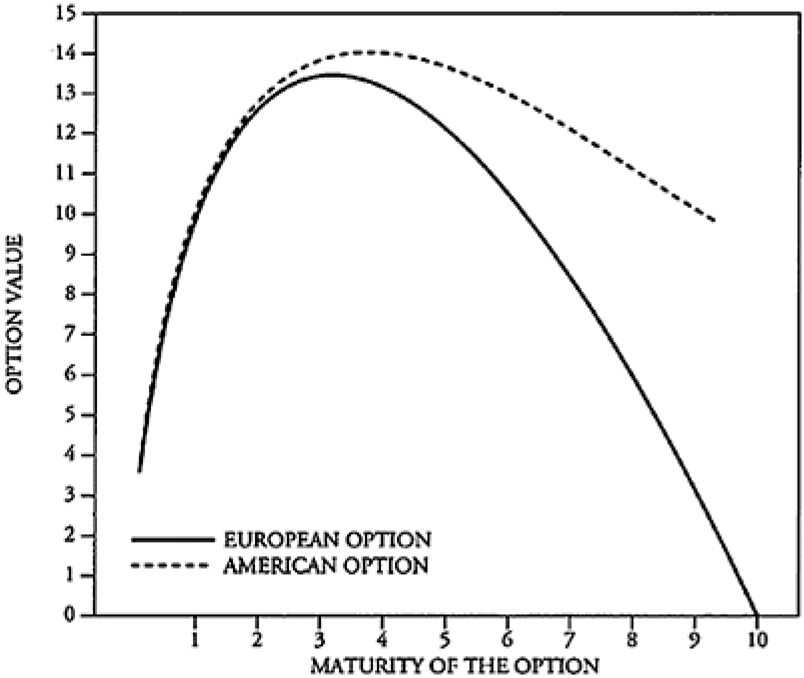

Figure 5.2 European and American call options on discount bonds

This figure shows the value of a European and American call option on a discount bond with face value 100 for different maturities of the option using the Merton model. The maturity of the underlying bond is ten years and the exercise price of the options is equal to the forward price of the underlying bond. The instantaneous short rate is 0.07, while the risk-neutral drift term and volatility are 0.02 and 0.07, respectively.

the time value of the claim. Finally, the face value of the bond is 100. As an interesting comparison to options on stocks, the figure also shows the values of the corresponding American call options.[2]

When the maturity of the option increases in comparison to the maturity of the underlying bond, two different effects occur. On the one hand, the increased maturity of the claim results in higher in-the-money option values at maturity and, therefore, in a corresponding higher current option value. On the other hand, as the underlying bond nears maturity, the range of possible bond values narrows, which decreases the current value of the at-the-money option. This second effect is partly offset if American options are considered. Although near the maturity of the bond a high option value is still less likely, the American feature of the option enables a premature exercise when high bond values occur. The option value, however, also decreases in this case, as maturity increases because of the higher exercise price that has to be paid.

To investigate the impact of volatility on option prices, Figure 5.3 shows the values of European at-the-money call options on discount bonds for different maturities and different volatilities. The parameter values and claim characteristics in this example are the same as in the previous figure, demonstrating the impact of volatility on the term structure of interest rates.

As volatility increases, the value of the at-the-money option generally increases and the two countervailing effects just described are more pronounced.

The Vasicek Model

The process of the instantaneous short rate in the Vasicek (1977) model obeys the following stochastic differential equation

$$dr(t) = \kappa(\theta - r(t))dt + \sigma dW(t) \tag{5.12}$$

The parameters *K*, θ and σ in this equation are real-valued constants, which are assumed to be positive, and $W=\{W(t), t \geq 0\}$ is a Standard Brownian Motion. The continuous time uncertainty is specified as in the Merton (1973) model, discussed in the previous section.

The drift term of this stochastic process forces the short-term rate of interest towards the parameter θ. If, for example, the value of $r(t)$ is greater than θ, the drift term is negative, which increases the probability of lower rates in the future. The degree with which this force is pulling back interest rates is dependent on κ, the mean reversion parameter. The distribution of the short rate at time t, given the information at time s, $s \leq t$, can again be obtained by making use of Arnold (1974):

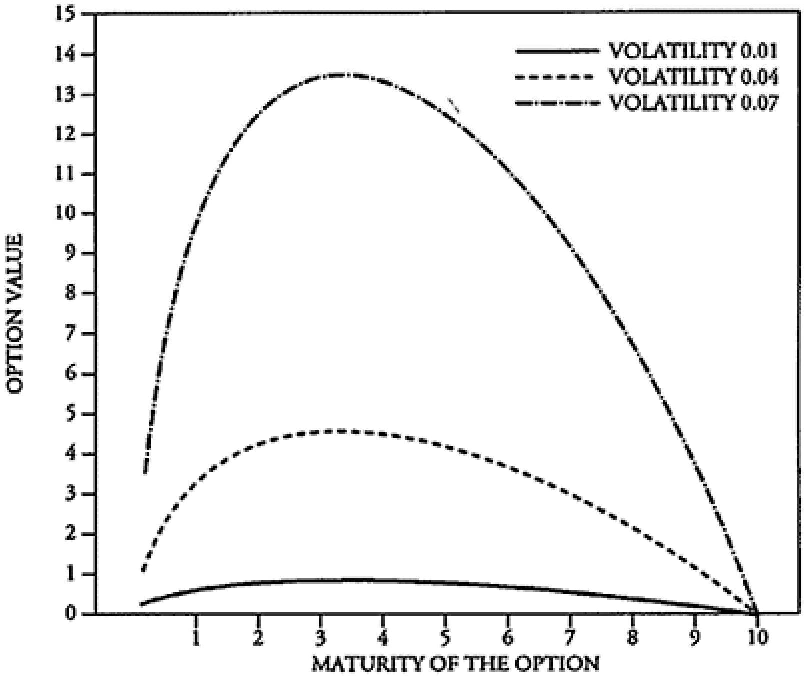

Figure 5.3 European call options on discount bonds

This figure shows the value of a European call option on a discount bond with face value 100 for different values of the volatility and for different maturities of the option in case of the Merton model. The maturity of the underlying bond is ten years and the exercise price of the options is equal to the forward price of the underlying bond. The instantaneous short rate is 0.07, while the risk-neutral drift term equals 0.02, respectively. The different values of the volatility are 0.01, 0.04 and 0.07, respectively.

$$r(t) \mid \mathcal{F}_s \sim N\left(\theta + (r(t) - \theta)e^{-\kappa(t-s)}, \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right)\right) \quad (5.13)$$

As in the Merton (1973) model, positive probabilities are assigned to negative interest rates. However, the unconditional mean and unconditional variance of the short rate at time *s*, given by

$$\lim_{t\to\infty} E_P(r(t) \mid \mathcal{F}_s) = \theta \tag{5.14}$$

$$\lim_{t\to\infty} Var_P(r(t) \mid \mathcal{F}_s) = \frac{\sigma^2}{2\kappa} \tag{5.15}$$

respectively, are both constant. The mean reversion process, therefore, implies significantly lower probabilities to unreasonably large interest rates and, as such, seems more suitable as a realistic description.

The existence and calculation of arbitrage-free values of interest rate derivative securities requires the existence of a unique equivalent probability measure such that values of trading strategies in terms of the money market account are martingales. The derivation of such an equivalent measure in this particular interest rate economy, starts again with the specification of the familiar Radon-Nikodym derivative $p=\{p(\lambda, t), t), t \geq 0\}$ as

$$\rho(\lambda, t) = exp\left(\int_0^t \lambda dW(s) - \frac{1}{2}\int_0^t \lambda^2 ds\right) \tag{5.16}$$

with the market price of risk λ. being a fixed real-valued constant. As before, Girsanov's Theorem can be applied to let the process $\tilde{W} = \{\tilde{W}(t), t \geq 0\}$, defined by

$$\tilde{W}(t) = W(t) - \int_0^t \lambda ds \tag{5.17}$$

be a Standard Brownian Motion on the probability space $(\Omega, \mathcal{F}, F, Q)$The unique equivalent measure Qon this probability space is again given by

$$dQ = \rho(\lambda, t)dP \tag{5.18}$$

The value of a discount bond *P(r(t), t, τ)* is derived by taking the discounted expected value of this bond with respect to the measure Q,

$$P(r(t), t, \tau) = E_Q\left(e^{-\int_t^{t+\tau} r(s)ds} \mid \mathcal{F}_t\right) = \tag{5.19}$$

$$exp\left(\frac{1}{\kappa}(1 - e^{-\kappa\tau})(R(t,\infty) - r(t)) - \tau R(t,\infty) - \frac{\sigma^2}{4\kappa^3}(1 - e^{-\kappa\tau})^2\right)$$

In this expression, $R(t, \infty)$ denotes the yield on a bond with infinite maturity, which is equal to the constant,

$$R(t,\infty)=\lim_{\tau\to\infty}R(t,\tau)=\theta+\lambda\frac{\sigma}{\kappa}-\frac{1}{2}\frac{\sigma^2}{\kappa^2} \qquad (5.20)$$

The instantaneous rate of return and volatility of this return on a discount bond are equal to $r(t)+\lambda\sigma(1-e^{-k\tau})/\kappa$ and $\sigma(1-e^{-k\tau})/\kappa$, respectively, under the original measure *P*. As opposed to the Merton (1973) model, where the return and volatility depend linearly on maturity, the instantaneous return and volatility in the Vasicek (1977) model increase nonlinearly, as maturity decreases with limiting values $r(t)+\lambda\sigma/\kappa$ and σ/κ, respectively, assuming a positive market price of risk.

The term structure of interest rates at time *t* is now,

$$R(t,\tau)=R(t,\infty)+\frac{1-e^{-\kappa\tau}}{\kappa\tau}(r(t)-R(t,\infty))+\frac{\sigma^2}{4\kappa^3\tau}(1-e^{-\kappa\tau})^2 \qquad (5.21)$$

It is interesting to note that the yield on a discount bond is a weighted combination of the instantaneous spot rate and the "infinite" maturity yield, with positive weights summing to one plus some function. This function is a product of spot rate volatility and a positive weighting function, which is zero at both ends of the maturity spectrum. As a result, this function causes a curvature of the yield curve. Figure 5.4 shows the term structure of interest rates for different values of the mean reversion parameter *K*. The initial value of the instantaneous spot rate is 0.07, while the values of the risk-neutral unconditional mean $\theta+\lambda\frac{\sigma}{\kappa}$ and volatility are 0.10 and 0.04, respectively. The different values of the mean reversion parameter are 0.2, 0.4 and 0.6.

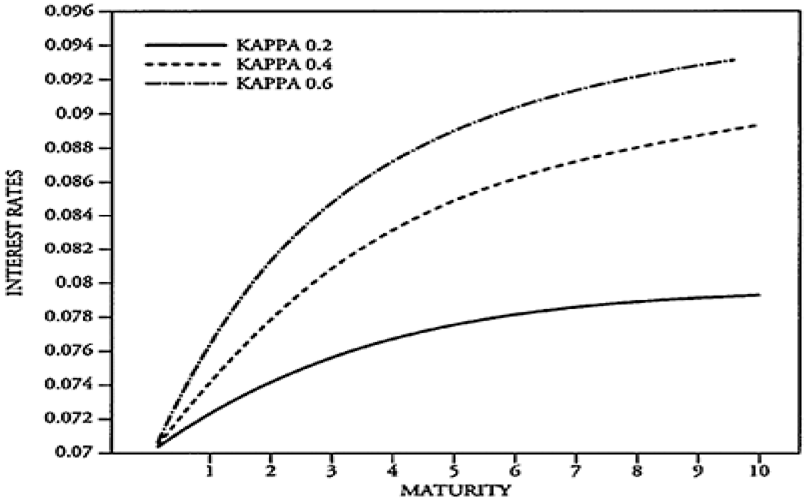

Figure 5.4 The term structure of interest rates

This figure shows the term structure of interest rates according to the model of Vasicek for different values of *K*. The instantaneous short-term rate of interest is 0.07, while the values of the risk-neutral unconditional mean and volatility are 0.10 and 0.04, respectively. The different values of the mean reversion parameter are 0.2, 0.4, and 0.6.

The different positions of the three term structures can simply be explained by looking at the "infinite" maturity yield and the weights of this yield and the short rate in the term structure expression. The higher the value of κ, the higher the value of the "infinite" maturity yield, thus rotating the yield curve. In addition, an increased value of κ puts more weight on the "infinite" maturity yield in comparison to the instantaneous short-term rate of interest, thereby increasing the curvature of the term structure. It is further interesting to note that the "infinite" maturity yield is always lower than the risk-neutral unconditional mean of the instantaneous spot rate. As is explained in the previous section, this is due to the convex relation between bond values and spot rates.

The distribution of a spot rate at time t, given the information at time s, $t \geq s$, can easily be derived by using the linear relationship between spot rates and the instantaneous short rate,

$$R(t, \tau) \mid \mathcal{F}_s \sim N(\mu_R(s, t), \sigma_R^2(s, t)) \qquad (5.22)$$

with

$$\mu_R(s, t) = (1 - e^{-\kappa(t-s)}) \lim_{t \to \infty} E_P(R(t, \tau) \mid \mathcal{F}_s) + e^{-\kappa(t-s)} R(s, \tau)$$

$$\lim_{t \to \infty} E_P(R(t, \tau) \mid \mathcal{F}_s) = R(t, \infty) + \frac{1 - e^{-\kappa\tau}}{\kappa\tau}(\theta - R(t, \infty)) +$$

$$\frac{\sigma^2}{4\kappa^3\tau}(1 - e^{-\kappa\tau})^2$$

and

$$\sigma_R^2(s, t) = \left(\frac{1 - e^{-\kappa\tau}}{\kappa\tau}\right)^2 (1 - e^{-2\kappa(t-s)}) \frac{\sigma^2}{2\kappa}$$

The mean of the spot rate is a weighted sum of the current value and the unconditional mean of the spot rate. The endogenous term structure of interest rate volatilities $\sigma_R(S, t)$ is a decreasing function of time-to-maturity with limiting value zero. The mean reversion process of the short rate, therefore, implies short-term interest rates are less volatile than are longterm interest rates. Figure 5.5 shows the endogenous term structure of interest rate volatilities for different values of κ. The values of the parameters used in this example are the same as in the previous figure. The difference *t–s* is equal to one year. The higher the value of κ, the lower the volatility of spot rates. As κ increases, the force with which interest rates are pulled back to their long-term mean increases, too. As a result, the volatility of corresponding interest rates decreases.

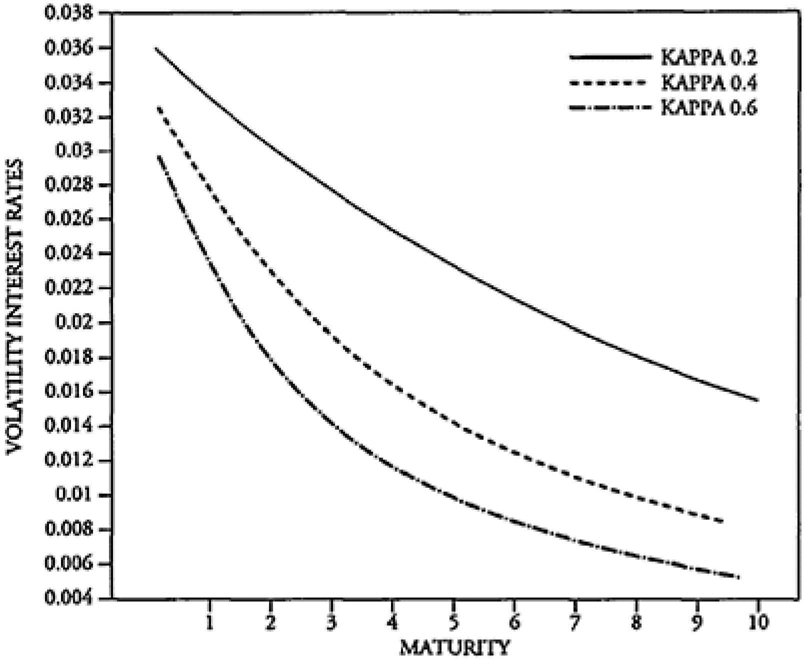

Figure 5.5 The term structure of interest rate volatilities

This figure shows the term structure of interest rate volatilities according to the model of Vasicek for different values of κ. The instantaneous short-term rate of interest is 0.07, while the values of the risk-neutral unconditional mean and volatility are 0.10 and 0.04, respectively. The different values of the mean reversion parameter are 0.2, 0.4, and 0.6. The difference *t–s* is equal to one year.

The value of a European call option on a discount bond is obtained by taking the familiar discounted expectation with respect to Q of the payout of the claim at maturity,

$$E_Q\left(e^{-\int_t^{t+\tau_1} r(s)ds} Max(P(r(\tau_1), \tau_1, \tau_2 - \tau_1) - K, 0)\right) \tag{5.23}$$

which is equal to

$$C(r(t), t, K, \tau_1, \tau_2) = P(r(t), t, \tau_2)N(d_1) - KP(r(t), t, \tau_1)N(d_2) \tag{5.24}$$

with

$$d_1 = \frac{1}{v}\ln\left(\frac{P(r(t), t, \tau_2)}{KP(r(t), t, \tau_1)}\right) + \frac{1}{2}v$$

$$d_2 = d_1 - v$$

$$v^2 = \frac{1}{2}\frac{\sigma^2}{\kappa^3}\left((1 - e^{-\kappa(\tau_2-\tau_1)})^2 - (e^{-\kappa(\tau_2-t)} - e^{-\kappa(\tau_1-t)})^2\right)$$

To assess the impact of the mean reversion parameter κ on option values, Figure 5.6 presents these values for different values of κ and different time-to-maturities of the option. The values of the other parameters are the same as in the previous examples. The maturity and face value of the underlying bond are ten years and 100, respectively, while the exercise

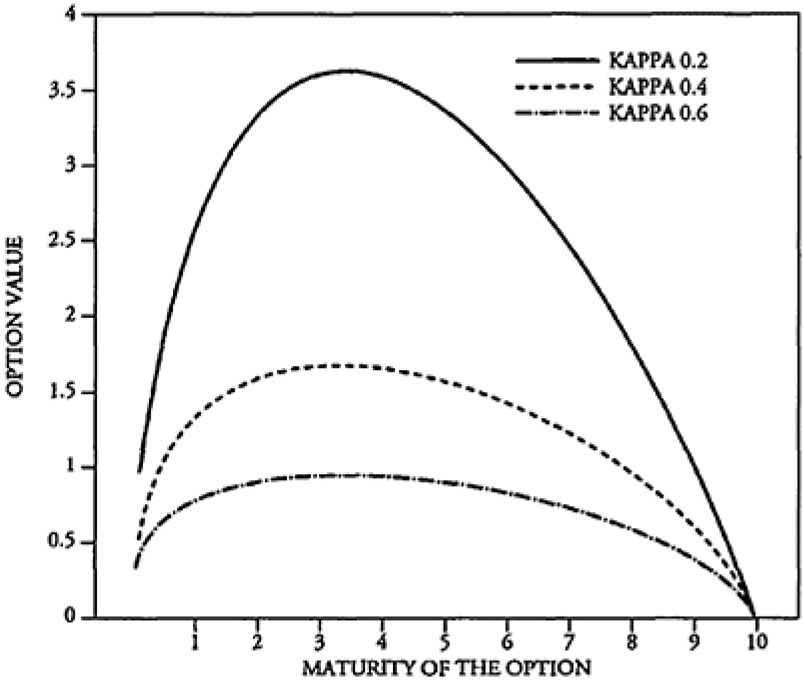

Figure 5.6 European call options on discount bonds

This figure shows the value of a European call option on a discount bond with face value 100 for different values of κ in the case of the Vasicek model. The maturity of the underlying bond is ten years and the exercise price of the option is equal to the forward price of this underlying bond. The instantaneous shortterm rate of interest is 0.07, while the values of the risk-neutral unconditional mean and volatility are 0.10 and 0.04, respectively. The different values of the mean reversion parameter are 0.2, 0.4, and 0.6.

price of the call option is equal to the forward price of this underlying bond to concentrate on the time value of the claim.

As expected, a higher value of κ yields a lower value of the call option. The two countervailing effects described in the previous section are less pronounced in this case due to the steeper term structure of interest rate volatilities.[3]

The Cox, Ingersoll and Ross Model

The basic drawbacks of the Merton (1973) model are the normality of interest rates and, related to the specific distribution chosen, the unbounded unconditional expectation and variance of the short rate. In the Vasicek (1977) model, this unboundedness of the first and second moment of the distribution of the instantaneous spot rate is avoided by assuming a mean reverting process of the short rate. The mean reversion is forcing the spot rate towards its unconditional mean, which is a constant. Although the first and second moment of the distribution of interest rates seem to be more reasonable in this model, interest rates are still normally distributed, implying positive probabilities of negative rates.

In Cox, Ingersoll and Ross (1985) (CIR), however, a stochastic process of the instantaneous short rate is proposed that retains the mean reversion property of the Vasicek model and excludes negative interest rates. In their model, the short rate obeys

$$dr(t) = \kappa(\theta - r(t))dt + \sigma\sqrt{(r(t))}dW(t) \tag{5.25}$$

The parameters κ, θ and σ are positive real-valued constants and the process $W = \{W(t), t \geq 0\}$ is a Standard Brownian Motion. The probability space and the corresponding probability measure are specified as in the previous two sections.

As in the Vasicek model, the drift term is causing a mean reversion of the spot rate. The interest rate $r(t)$ is stochastically moving around a location parameter θ with speed of adjustment or mean reversion force κ. If $2\kappa\theta \geq \sigma^2$, the upward drift is sufficiently large to make the origin inaccessible.[4] If initial interest rates are positive, they can never become negative with probability one.

The actual distribution function of spot rates at time t, given the information at time s, $s \leq t$, is the non-central chi-square, $x(2cr(t), 2q+2, 2u)$, with $2q+2$ degrees of freedom and parameter of non-centrality $2u$ proportional to the current spot rate $r(s)$. The constants c, u, v and q are defined as

$$c = \frac{2\kappa}{\sigma^2(1 - e^{-\kappa(t-s)})}$$

$$u = cr(s)e^{-\kappa(t-s)}$$

$$v = cr(t)$$

$$q = \frac{2\kappa\theta}{\sigma^2} - 1$$

According to this distribution, the mean and variance of the spot rate at time /, given the information at time *s*, $s \leq t$, are

$$E_P(r(t) \mid \mathcal{F}_s) = \theta + (r(t) - \theta)e^{-\kappa(t-s)} \tag{5.26}$$

$$Var_P(r(t) \mid \mathcal{F}_s) = r(s)\left(\frac{\sigma^2}{\kappa}\right)(e^{-\kappa(t-s)} - e^{-2\kappa(t-s)}) + \tag{5.27}$$

$$\theta\left(\frac{\sigma^2}{2\kappa}\right)(1 - e^{-\kappa(t-s)})^2$$

The mean is a weighted average of the current value of the spot rate and the unconditional mean *θ*. These weights are positive and sum to one, reflecting the mean reversion of spot rates. As opposed to the model of Vasicek, the variance is positively dependent both on the current value of the spot rate and on the unconditional mean. As *t* becomes larger, the variance becomes independent on the current information at time *s*. The steady-state mean and variance of the spot rate are, summarized,

$$\lim_{t\to\infty} E_P(r(t) \mid \mathcal{F}_s) = \theta \tag{5.28}$$

$$\lim_{t\to\infty} Var_P(r(t) \mid \mathcal{F}_s) = \theta\left(\frac{\sigma^2}{2\kappa}\right) \tag{5.29}$$

To derive the arbitrage-free values of interest rate dependent securities, such as bonds and option on bonds, the study proceeds along the lines of the previous sections. The first step in this derivation is to determine a unique probability measure, which is equivalent to *P*, by specifying the familiar Radon-Nikodym derivative p={P(λ, r(t), t), *r(t), t)*, $t \geq 0$} as

$$\rho(\lambda, r(t), t) = exp\left(\int_0^t \lambda\sqrt{r(s)}dW(s) - \frac{1}{2}\int_0^t \lambda^2 r(s)ds\right) \tag{5.30}$$

with A a fixed real-valued constant. In contrast to the models of Merton and Vasicek, the change of probability measure is explicitly dependent on the value of the short rate. As the remainder of this section will demonstrate, this particular choice is justified because of the resulting analytic tractability of the model. The characteristics of the stochastic differential equation of the short rate *r(t)* allows for application of Girsanov's Theorem. The process $\tilde{W} = \{\tilde{W}(t), t \geq 0\}$, defined by

$$\tilde{W}(t) = W(t) - \int_0^t \lambda\sqrt{r(s)}ds \qquad (5.31)$$

is then a Standard Brownian Motion on the probability space $(\Omega, \mathcal{F}, F, Q)$. The unique equivalent measure Q on this space is given by

$$dQ = \rho(\lambda, r(t), t)dP \qquad (5.32)$$

The value of a discount bond *P*(*r(t), t, τ*) is obtained by taking the discounted expected value of this bond with respect to the measure Q,

$$P(r(t), t, \tau) = E_Q\left(e^{-\int_t^{t+\tau} r(s)ds} \mid \mathcal{F}_t\right) = A(\tau)e^{-B(\tau)r(t)} \qquad (5.33)$$

with

$$A(\tau) = \left(\frac{2\gamma e^{(\tilde{\kappa}+\gamma)\tau/2}}{(\tilde{\kappa}+\gamma)(e^{\gamma\tau}-1)+2\gamma}\right)^{\frac{(\tilde{\kappa}+\gamma)R(t,\infty)}{\sigma^2}}$$

$$B(\tau) = \frac{2(e^{\gamma\tau}-1)}{(\tilde{\kappa}+\gamma)(e^{\gamma\tau}-1)+2\gamma}$$

$$\gamma = (\tilde{\kappa}^2 + 2\sigma^2)^{\frac{1}{2}}$$

and

$$\tilde{\kappa} = \kappa - \lambda\sigma$$

In this expression, *R*(*t*, ∞) denotes the "infinite" maturity yield, which is equal to the constant

$$R(t, \infty) = \lim_{\tau\to\infty} R(t, \tau) = \frac{2\kappa\theta}{\tilde{\kappa}+\gamma}$$

Under the original measure *P*, the instantaneous rate of return and volatility of this return on a discount bond are $r(t)(1+\lambda\sigma B(\tau))$ and $\sqrt{r(t)}\sigma B(\tau)$, respectively. Assuming a positive market price of risk $\lambda\sqrt{r(t)}$, an increase in the maturity of the bond increases this return and volatility monotonically, with limiting values $r(t)(1+2\lambda\sigma/(\tilde{\kappa}+\gamma))$ *and* $\sqrt{r(t)}\sigma/(\tilde{\kappa}+\gamma)$, respectively.

The term structure of interest rates at time / is equal to

$$R(t,\tau) = w_1(\tilde{\kappa},\sigma,\tau)r(t) + w_2(\tilde{\kappa},\sigma,\tau)R(t,\infty) \quad (5.34)$$

with

$$w_1(\tilde{\kappa},\sigma,\tau) = \frac{B(\tau)}{\tau}$$

and

$$w_2(\tilde{\kappa},\sigma,\tau) = -\frac{(\tilde{\kappa}+\gamma)}{2\kappa\theta}\frac{\ln(A(\tau))}{\tau}$$

The yield on a zero-coupon bond is a weighted combination of the instantaneous spot rate and the “infinite” maturity yield, where the weights are strictly positive. It is interesting, now, to make a comparison between the term structure of interest rates based on the Vasicek and CIR models. In both models, the initial value of the spot rate *r(t)*, the mean reversion parameter κ and the unconditional mean θ are equal to 0.07, 0.4 and 0.10, respectively. The volatility in the Vasicek model is 0.04, while in the CIR model it is $0.04/\sqrt{0.07}$ to equal the instantaneous volatilities of spot rates. The market price of risk A. is equal to zero in both models. Figure 5.7 shows the term structures of interest rates based on these parameters.

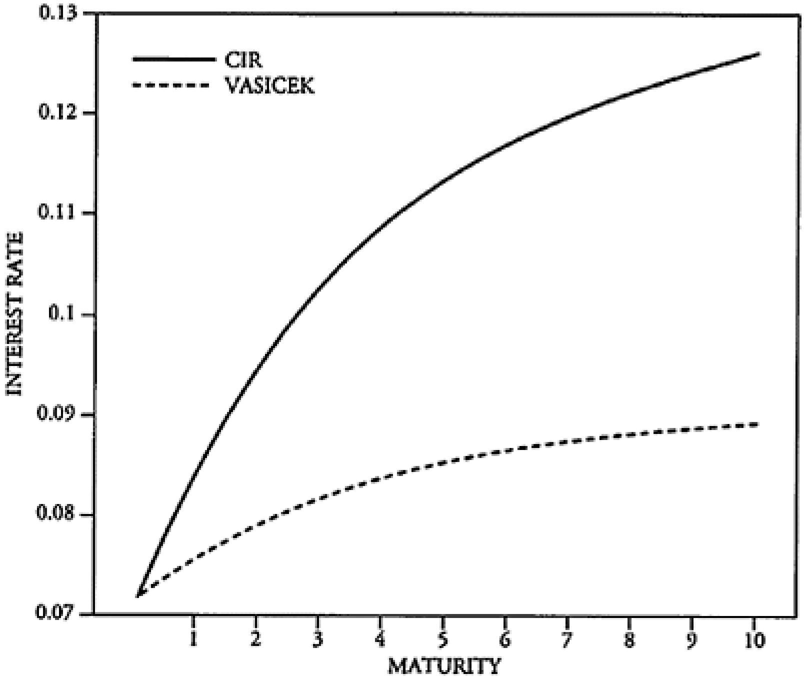

Figure 5.7 The term structure of interest rates

This figure shows the term structures of interest rates according to the models of Vasicek and Cox, Ingersoll and Ross (CIR). The instantaneous short-term rate of interest is 0.07, while the values of the mean reversion parameter and the unconditional mean are 0.4 and 0.10, respectively. The value of the volatility parameter using the Vasicek model is 0.04 and using the CIR model, $0.04/\sqrt{0.07}$.

Interest rates based on the CIR model are always lower than those based on the Vasicek model. Because the instantaneous volatility of the spot rate is linearly dependent on the square root of the spot rate in the CIR model, a lower value of this rate increases the probability of a relatively low rate one instant later. Although a similar argument can be posed for the case of increased probabilities of relatively high spot rates in the CIR model, the convex relation between bond values and spot rates results in significantly lower interest rates in comparison to the Vasicek model.

The linear relationship between interest rates and the instantaneous short-term rate of interest of equation (5.34), together with the analytic expressions of the mean and variance of the short rate (equations (5.26) and (5.27)) provide the means for obtaining the mean and variances of spot rates at time *t*, given the information at time *s*, $s \leq t$,

$$\mu_R(s,t)=(1-e^{-\kappa(t-s)})\lim_{t\to\infty}E_P(R(t,\tau)\mid\mathcal{F}_s)+e^{-\kappa(t-s)}R(s,\tau) \quad (5.35)$$

with

$$\lim_{t\to\infty}E_P(R(t,\tau)\mid\mathcal{F}_s)=w_1(\tilde{\kappa},\sigma,\tau)\theta+w_2(\tilde{\kappa},\sigma,\tau)R(t,\infty)$$

and

$$\sigma_R^2(s,t)=a_0(s,t,\tau)+a_1(s,t,\tau)R(s,\tau) \quad (5.36)$$

with

$$a_0(s,t,\tau)=\left(\frac{B(\tau)}{\tau}\right)^2\theta\left(\frac{\sigma^2}{2\kappa}\right)(1-e^{-2\kappa(t-s)})^2+$$

$$\left(\frac{B(\tau)}{\tau}\right)^2\left(\frac{\sigma^2}{\kappa}\right)(e^{-\kappa(t-s)}-e^{-2\kappa(t-s)})\frac{\ln(A(\tau))}{B(\tau)}$$

$$a_1(s,t,\tau)=\frac{B(\tau)}{\tau}\left(\frac{\sigma^2}{\kappa}\right)(e^{-\kappa(t-s)}-e^{-2\kappa(t-s)})$$

The variance of interest rates is a linear function of this interest rate. As time increases, the variance decreases monotonically with limiting value

$$\lim_{t\to\infty}Var_P(R(t,\tau)\mid\mathcal{F}_s)=\left(\frac{B(\tau)}{\tau}\right)^2\theta\left(\frac{\sigma^2}{2\kappa}\right) \quad (5.37)$$

As in the Vasicek model, longer-term interest rates are less volatile than short-term interest rates are.

Figure 5.8 shows the term structure of volatilities of the CIR and Vasicek models for the same set of parameters as the previous figure.

The volatility of interest rates in the CIR model is significantly lower than that in the Vasicek model. The volatility curve, in addition, is much

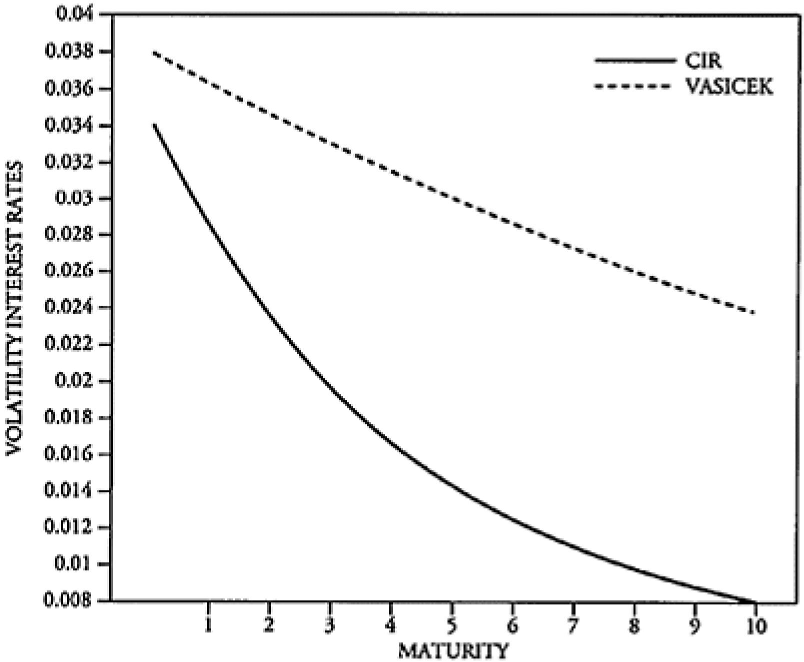

Figure 5.8 The term structure of interest rate volatilities

This figure shows the term structures of interest rate volatilities according to the models of Vasicek and Cox, Ingersoll and Ross (CIR). The instantaneous short-term rate of interest is 0.07, while the values of the mean reversion parameter and the unconditional mean are 0.4 and 0.10, respectively. The value of the volatility parameter using the Vasicek model is 0.04 and using the CIR model, $0.04/\sqrt{0.07}$.

steeper. Compare this to Figure 5.7, where the term structures of interest rates were shown for both models, and this can be explained similarly. The linear relationship between the variance of the spot rate and its value increases the probability of low rates in the future, when current rates are relatively low. Although this can also happen the other way around, the generally lower interest rate volatilities are due to the convex relation between bonds and instantaneous short rates. As the maturity of interest rates increases, this combined variance and convexity effect increases, too, resulting in a steeper term structure of interest rate volatility in comparison to the Vasicek model.

Next the Cox, Ingersoll and Ross model will be applied by discussing the derivation and characteristics of a European call option on a discount bond. To obtain this value, calculate the discounted expectation with respect to the probability measure Q of the payout of the claim atmaturity:

$$E_Q\left(e^{-\int_t^{t+\tau_1} r(s)ds} Max(P(r(\tau_1), \tau_1, \tau_2 - \tau_1) - K, 0)\right) \tag{5.38}$$

This yields (see CIR (1985, p. 396))

$$C(r(t), t, K, \tau_1, \tau_2) = P(r(t), t, \tau_2)\chi\left(d_1, \frac{4\kappa\theta}{\sigma^2}, \frac{\phi^2 r(t) e^{\gamma\tau_1}}{d_1/r^*}\right) \tag{5.39}$$

$$- KP(r(t), t, \tau_1)\chi\left(d_2, \frac{4\kappa\theta}{\sigma^2}, \frac{\phi^2 r(t) e^{\gamma\tau_1}}{d_2/r^*}\right)$$

with

$$d_1 = 2r^*(\phi + \psi + B(\tau_2 - \tau_1))$$

$$d_2 = 2r^*(\phi + \psi)$$

$$\phi = \frac{2\gamma}{\sigma^2(e^{\gamma\tau_1} - 1)}$$

$$\psi = \frac{\tilde{\kappa} + \gamma}{\sigma^2}$$

and

$$r^* = \ln\left(\frac{A(\tau_2 - \tau_1)}{K}\right) \Big/ B(\tau_2 - \tau_1)$$

which is the value of the instantaneous spot rate at maturity of the option at which investors are indifferent between exercising the option or not.

Although not obvious at first sight, this option pricing formula has an interpretation similar to those obtained in previous sections. The first term in the option formula is the discounted expected value of the bond on which the option is written, conditional upon an in-the-money option at maturity of the option. The second term is the discounted exercise price of the option multiplied by the probability of the option ending up in-the-money.

Figure 5.9 compares option prices based on the Vasicek and CIR model, for different time-to-maturities. The parameters are again the same as in Figures 5.7 and 5.8. The option is a European at-the-money call option on a discount bond with a face value of 100 and a time-to-maturity equal to ten years.

It would be logical to expect the option values based on the CIR model to be lower than those based on the Vasicek model. As options depend crucially on the term structure of volatilities, the previous figure, in which both models were compared with respect to these term structures, would suggest such a result. However, the CIR option values are higher for

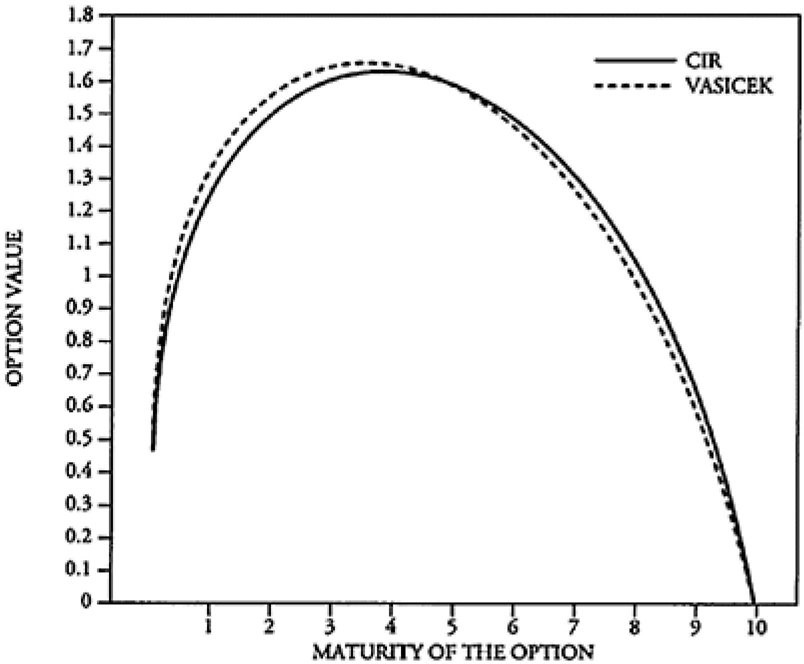

Figure 5.9 European call options on discount bonds

This figure shows the value of a European call option on a discount bond with face value 100 for different time-to-maturities using the models of Vasicek and Cox, Ingersoll and Ross (CIR). The exercise price of the option is equal to the forward price of the underlying bond. The instantaneous short-term rate of interest is 0.07, while the values of the mean reversion parameter and the unconditional mean are 0.4 and 0.10, respectively. The value of the volatility parameter using the Vasicek model is 0.04 and using the CIR model, $0.04/\sqrt{0.07}$.

most of the maturity range considered. This result is due to the particular shapes of the term structures of volatilities.

The discussion of the option pricing formulas made it clear that European option prices are a function of the volatility of the forward price of the underlying discount bond. As the term structure of volatilities in the CIR model is, for most of the maturity spectrum, much steeper than in the Vasicek model regardless of the volatilities' specific position, forward price volatilities and corresponding option prices are higher.[5]

Miscellaneous Interest Rates Models

The previous section discussed some well-known interest rates models. These models have the common property that they allow for closed-form solutions for the term structure of interest rates, the term structure of interest rate volatilities and European call options and put options on discount bonds. Connected to this analytic tractability, a great deal of attention has been paid to investigating these models empirically. As will be seen in later chapters, the model of Cox, Ingersoll and Ross (1985), especially, has been the subject of a thorough empirical examination.

This section will briefly discuss some other interest rates models. The basic drawback of these models is the necessity of using numerical methods to obtain, for example, the term structure of interest rates. This dependence limits an economic interpretation and practical implementation of these models. As the main reason for discussing these models is to complete the overview of the endogenous term structure of interest rates models, no attention will be given to a clear description of probability spaces and probability measures.

In the interest rates model of Dothan (1978), the instantaneous rate of interest obeys the following stochastic differential equation

$$dr(t) = \sigma r(t) dW(t) \tag{5.40}$$

In this specification, *a* is a real-valued constant and $W=\{W(t),\ t \geq 0\}$ is a familiar Standard Brownian Motion. The zero drift term and proportional volatility function cause the spot rate $r(t)$ at time t, given the infomation at time s, $s \leq t$, to be lognormally distributed with mean $r(s)$ and variance $r(s)^2(e^{\sigma^2(t-s)}-1)$and, more specifically, to be positive with probability one.[6]

As in the previous illustrations, a unique equivalent probability measure can be determined under which relative values of interest rate derivative securities are martingales. The resulting term structure of interest rates is a monotonically decreasing function of maturity, an increasing concave function of the current value of the instantaneous spot rate and a decreasing convex function of σ^2. The analytic expression, however, of this term structure of interest rates is quite complicated, making a statistical analysis cumbersome. The valuation of other interest rate derivative securities, such as options on bonds, can be performed only by using numerical methods.[7]

In the model of Courtadon (1982), the stochastic process of the instantaneous short rate of the Dothan model is extended by a mean reverting drift term, that is,

$$dr(t) = \kappa(\theta - r(t))dt + \sigma r(t)dW(t) \tag{5.41}$$

The speed of adjustment parameter κ, the unconditional mean of interest rates θ and the volatility parameter σ are real-valued constants and $W = \{W(t), t \geq 0\}$ is again a Standard Brownian Motion. Although the distribution of spot rates in this model is unknown and no analytic expression can be obtained for the moments of the distribution, it is argued in Courtadon that this process is more suited to describe the actual behavior of interest rates than is the Geometric Brownian Motion because interest rates exhibit mean reversion and are strictly positive with probability one.[8] As this may be true, it is by no means certain, however, that this process is to be preferred empirically to the mean reverting processes of Vasicek or Cox et al. apart from the analytic tractability of these models.

After the derivation of a unique equivalent probability measure, values of interest rate derivative securities can only be obtained by use of numerical methods.

In the model of Brennan and Schwartz (1983), interest rate derivative securities are assumed to be a function of the instantaneous short-term rate of interest and the yield on a consol bond or long rate. These short-rate and long-rate processes $r=\{r(t), t \geq 0\}$ and /=$\{l(t), t \geq 0\}$ obey the following stochastic differential equations,

$$dr(t) = \theta_r dt + \sigma_{r1} dW_1(t) + \sigma_{r2} dW_2(t)$$

$$dl(t) = \theta_l dt + \sigma_{l1} dW_1(t) + \sigma_{l2} dW_2(t)$$

The parameters θ_r, θ_l, σ_{r1}, σ_{r2}, σ_{l1} and σ_{l2} are real-valued constants and W“= $\{W_1(t), W_2(t), t \geq 0\}$ is a two-dimensional vector process of independent Standard Brownian Motions.

The exclusion of arbitrage opportunities results in a second order partial differential equation, which has to be solved numerically for all interest rate derivative securities, making use of particular boundary conditions.

As the value of a consol bond at time t is a closed-form solution[9] of the consol rate /(/), interest rate securities are a function only of the risk-neutral drift term of the short rate $\tilde{\theta}_r$ and the different volatility parameters σ_{r1}, σ_{r2}, σ_{l1} *and* σ_{l2} However, because some of the coefficients in this partial differential equation are unbounded in the underlying stochastic processes, the numerical results obtained are questionable.

EXOGENOUS TERM STRUCTURE OF INTEREST RATES MODELS

Within the class of endogenous term structure of interest rates models, all securities are considered to be interest rate contingent claims. Given the distributional characteristics of the instantaneous short rate, the exclusion of arbitrage opportunities results in valuation relationships for any interest rate derivative security. As an important result, the term structure of interest rates at some valuation date is a specific function of the current value of the short rate and its risk-neutral stochastic characteristics. Thus, not only the stochastic movement over time of the yield curve is implied by the short rate, but also the actual shape of the yield curve at the valuation date.

Within the class of exogenous term structure of interest rates models, however, the stochastic differential equation of the short rate is specified such that a given term structure of interest rates or term structure of forward rates at an initial valuation date is obtained. Instead of implying the actual shape of the yield curve and its subsequent stochastic movement over time, the stochastic component of the process of the short rate now determines only the distribution of interest rates.

The model of Ho and Lee (1986) is the first model that is based on this fundamentally different approach of valuing interest rate contingent claims. Given the prices of all zero-coupon bonds at the valuation date, a binomial lattice is constructed representing the discrete stochastic movement of these bond prices or term structures over time. In Heath *et al.* (1990a, 1992), this approach is considerably extended. Besides the generalization to continuous time, a whole family of potential stochastic processes is presented to describe the movement of the forward rate curve over time, and general conditions are derived to exclude riskless arbitrage opportunities.

The different models implying a given term structure of interest rates at the valuation date can be further classified according to the overview of Chapter 3. In the first part of this section, those models are presented in which the stochastic differential equation of the instantaneous short rate implies a term structure of interest rate volatilities. The second part of this section will then discuss an approach to modify the stochastic diflferential equation to reflect both a given term structure of interest rates as well as a given term structure of interest rate volatilities.

The presentation and discussion of the different models in this section is completely opposite to that in the previous section, starting with the derivation of the stochastic differential equation of the short-term rate of interest in a risk-neutral economy such that the initial yield curve at the valuation date is obtained and arbitrage opportunities do not exist. Subsequently, the possible term structures of interest rates after the initial trading date will be obtained and, if possible, the closed-form value of a European call option on a discount bond will be derived. As investors do not have to be risk-neutral and may have subjective probability beliefs with respect to a particular assessment of certain events, the actual distribution of interest rates and the endogenous and exogenous term structure of interest rate volatilities will be derived. For the same reasons as in the previous section, limited attention is paid to the multi-factor analogues of the one-factor models discussed in this section.

Endogenous Term Structure of Volatility Models

This subsection begins with the discussion of the continuous time version of the Ho and Lee (1986) model. Because the Ho and Lee model is a

counterpart to the Merton (1973) model, attention can be focused on the differences between the endogenous and exogenous term structure of interest rates models. An extension of this will be provided by examining another model within the class of Heath *et al.* (1992) term structure models, which is a counterpart to the Vasicek (1977) model discussed in the previous section. After this, the general model of Heath *et al.* will be examined.

The Ho and Lee Model

The continuous time uncertainty of the risk-neutral economy is specified by the filtered probability space $(\Omega, \mathcal{F}, F, Q)$, satisfying the usual conditions and a trading interval [0, *T*] for a fixed $T > 0$. As the economy is assumed to be risk-neutral, investors have to agree on the unique probability measure Q in this particular type of economy. The exogenous term structure of interest rates at the initial trading date 0 is given and represented by the initial term structure of forward rates $f(0, t)$, t [0, *T*], which is non-random, measurable and absolutely integrable and defined as

$$P(0, \tau) = e^{-\int_0^\tau f(0,s)ds} \tag{5.42}$$

for all τ [0, *T*].

The value of the instantaneous short-term rate of interest at time t [0, *T*], is assumed to obey

$$r(t) = \alpha(t) + \sigma \tilde{W}(t) \tag{5.43}$$

In this expression, $\alpha(t)$, t [0, *T*] is non-random, measurable and absolutely integrable, a is a real-valued constant and $\tilde{W} = \{\tilde{W}(t), t \in [0, T]\}$ is a Standard Brownian Motion, initialized at zero.[10]

The exclusion of arbitrage opportunities requires the existence of a unique equivalent probability measure such that the values of all interest rate derivative securities relative to the money market account are martingales. By definition, this probability measure is Q, and by imposing suitable conditions with respect to the function $\alpha(t)$, t [0, *T*], we can ensure that the given term structure of forward rates at the initial trading date 0 is obtained. The value of a discount bond P(0, τ) at time 0 with maturity τ, τ [0, *T*], therefore, has to be equal to the following two expressions,

$$P(0, \tau) = e^{-\int_0^\tau f(0,s)ds} = E_Q\left[e^{-\int_0^\tau r(s)ds} \mid \mathcal{F}_0\right] \tag{5.44}$$

Calculating the expectation,[11] results in

$$\alpha(t) = f(0, t) + \frac{1}{2}\sigma^2 t^2 \qquad (5.45)$$

for all t [0, T].

To obtain the value of a discount bond $P(t, \tau)$ at some time t [0, T] with maturity τ, the discounted expected value is calculated with respect to the unique probability measure Q,

$$P(t, \tau) = E_Q\left[e^{-\int_t^{t+\tau} r(s)ds} \mid \mathcal{F}_t\right] \qquad (5.46)$$

$$= exp\left(-\int_t^{t+\tau} f(0, s)ds - \tau\sigma\tilde{W}(t) - \frac{1}{2}\sigma^2 t\tau(t+\tau)\right)$$

The possible term structures of interest rates at some time t [0, T] can be derived from the values of these bonds, as they are implied by the term structure of interest rates at the initial trading date and the stochastic movement of the short rate over time,

$$R(t, \tau) = -\frac{\ln P(t, \tau)}{\tau} = \frac{\int_t^{t+\tau} f(0, s)ds}{\tau} + \sigma\tilde{W}(t) + \sigma^2 t(t+\tau) \qquad (5.47)$$

Using relationship (5.43), this can be rewritten as

$$R(t, \tau) = r(t) - f(0, t) + \frac{\int_t^{t+\tau} f(0, s)ds}{\tau} + \frac{1}{2}\sigma^2 t\tau \qquad (5.48)$$

This final expression, in which future yields are related to the value of the short rate, resembles the similarity between the models of Merton and Ho and Lee. The value of a particular yield at time t [0, T] with maturity τ is the sum of the difference between the instantaneous short rate and forward rate, the forward rate at the initial trade 0 covering the period $[t, t+\tau]$, and a term reflecting the convexity between interest rates and bond prices. As in the Merton model, changes in the short rate cause parallel shifts in the term structure. However, although the continuous time economy is not designed to look at infinite maturity yields due to the fixed final trading date $T > 0$, the infinite maturity yield at some future date t [0, T] in the Ho and Lee model has limiting value if forward rates are bounded

$$\lim_{\tau \to \infty} R(t, \tau) = \infty \qquad (5.49)$$

Because the risk-neutral drift term of the short rate increases without bound as maturity increases, the convexity effect causing the infinite maturity yield to decrease without bound in the Merton model is completely outweighted.[12]

The value of a European call option $C(K, \tau_1, \tau_2)$ at the initial valuation date with maturity τ_1 written on a discount bond with maturity τ_2, is derived analogously,

$$C(K,\tau_1,\tau_2)=E_Q\left(e^{-\int_0^{\tau_1} r(s)ds}Max(P(\tau_1,\tau_2-\tau_1)-K,0)\right) \quad (5.50)$$

and equal to

$$C(K,\tau_1,\tau_2)=P(0,\tau_2)N(d_1)-KP(0,\tau_1)N(d_2) \quad (5.51)$$

with

$$d_1=\frac{1}{v}\ln\left(\frac{P(0,\tau_2)}{KP(0,\tau_1)}\right)+\frac{1}{2}v$$

$$d_2=d_1-v$$

$$v^2=\sigma^2\tau_1(\tau_2-\tau_1)^2$$

Besides the actual values of the short and long bond at the valuation date, this option pricing formula is equal to the one derived and already illustrated in Section 5.1 where the Merton model was discussed.

The possible bond values at some future date and European call option prices on a discount bond are derived by taking the appropriate expectations with respect to the probability measure Q. Investors agree on the existence of this unique measure in order to exclude arbitrage opportunities. However, the actual probability measure that implies the distribution of the term structure of interest rates may differ among investors, as was seen in the previous section where the endogenous term structure of interest rates models was examined. To derive the subjective distributions of interest rates, a stochastic process or Radon-Nikodym derivative $p=\{p(\lambda, t), t \quad [0, T]\}$ is defined by

$$\rho(\lambda,t)=exp\left(\int_0^t \lambda d\tilde{W}(s)-\frac{1}{2}\int_0^t \lambda^2 ds\right) \quad (5.52)$$

In this case, the market price of risk, λ, is assumed to be a fixed, realvalued constant although other specifications are possible.[13] By Girsanov's Theorem, the stochastic process $W=\{W(t), t \quad [0, T]\}$, defined by

$$W(t)=\tilde{W}(t)-\int_0^t \lambda ds \quad (5.53)$$

is a Standard Brownian Motion on the probability space $(\Omega,\mathcal{F},F,P)$, where the equivalent probability measure P is given by

$$dP=\rho(\lambda,t)dQ \quad (5.54)$$

Given a fixed value of the market price of risk, λ, interest rates at time t [0, T], given the information at the initial trade date, are normally distributed[14] with respect to the probability measure P,

$$R(t,\tau)\mid\mathcal{F}_0\sim N\left(\frac{\int_t^{t+\tau}f(0,s)ds}{\tau}+\sigma\lambda t+\sigma^2 t(t+\tau),\sigma^2 t\right) \tag{5.55}$$

As investors may have different beliefs about the value of the market price of risk, λ, probability measures and corresponding distributions of interest rates may differ with respect to their mean. However, given the previous assumptions regarding the initial term structure of forward rates, the mean of future yields becomes unreasonably large as time increases, irrespective of the value of λ. As in the Merton model, the endogenous term structure of interest rate volatilities is again a flat function, independent of any particular maturity.

The Heath, Jarrow and Morton I Model

The basic drawbacks of the model of Ho and Lee (1986) are the positive probability of negative interest rates in the future and the unreasonable values of the corresponding mean of these future interest rates. As these problems are also present in the Merton (1986) model and solved partially in the Vasicek (1977) model by introducing mean reversion of the instantaneous short rate, it is a natural step to apply this approach to the Ho and Lee model. This analogue of the Vasicek model is discussed as an example in Heath *et al.* (1990a, 1992) and proposed after an empirical analysis by Dybvig (1989).

The exposition of this model starts with the specification of the continuous time uncertainty of the risk-neutral economy by the filtered probability space $(\Omega,\mathcal{F},F,Q)$, satisfying the usual conditions. There is a trading interval of [0, T], $T > 0$ and, because the economy is assumed to be riskneutral, all investors have to agree on a unique probability measure Q. The exogenous term structure of forward rates at the initial trade 0 is given by the function $f(0, t)$, t [0, T], which is again non-random, measurable and absolutely integrable.

The value of the instantaneous short-term rate of interest at time t [0, T] is assumed to obey[15]

$$r(t)=\alpha(t)+\sigma x(t) \tag{5.56}$$

The function $\alpha(t)$, t [0, T], is non-random, measurable and absolutely integrable, σ is a real-valued constant and $x(t)$, t [0, T] is the unique solution to the Ornstein-Uhlenbeck process

$$x(t)=\int_0^t e^{-\frac{\kappa}{2}(t-s)}d\tilde{W}(s) \tag{5.57}$$

where K is a real-valued constant and $\tilde{W} = \{\tilde{W}(t), t \in [0, T]\}$ is a Standard Brownian Motion, initialized at zero.

As before, the exclusion of arbitrage opportunities requires relative values of interest rate derivative securities to be martingales with respect to a unique equivalent probability measure. This measure is assumed to exist and already defined as Q. To ensure that the given term structure of interest rates at the valuation date is obtained, suitable restrictions upon the function $\alpha(t)$, t [0, T] have to be imposed. The value of a discount bond $P(0, \tau)$ at time 0 with maturity τ [0, T], therefore, has to be equal to the following two expressions:

$$P(0, \tau) = e^{-\int_0^\tau f(0,s)ds} = E_Q\left(e^{-\int_0^\tau r(s)ds} \mid \mathcal{F}_0\right) \tag{5.58}$$

Solving this expression[16] yields,

$$\alpha(t) = f(0, t) + \sigma^2 \frac{2}{\kappa^2}(1 - e^{-\frac{\kappa}{2}t})^2 \tag{5.59}$$

for all t [0, T].

The value of a discount bond $P(t, \tau)$ at some time t [0, T] in the future, is calculated in the same way,

$$P(t, \tau) = E_Q\left(e^{-\int_t^{t+\tau} r(s)ds} \mid \mathcal{F}_t\right) \tag{5.60}$$

yielding

$$exp\left(-\int_t^{t+\tau} f(0, s)ds - \sigma\frac{2}{\kappa}(1 - e^{-\frac{\kappa}{2}\tau})x(t) + v_1(t, \tau)\right) \tag{5.61}$$

with

$$v_1(t, \tau) = 2\frac{\sigma^2}{\kappa^3}(1 + (2 - e^{-\frac{\kappa}{2}(t+\tau)})^2 - (2 - e^{-\frac{\kappa}{2}\tau})^2 - (2 - e^{-\frac{\kappa}{2}t})^2) \tag{5.62}$$

The term structure of interest rates at time t [0, T] equals,

$$R(t, \tau) = -\frac{\ln P(t, \tau)}{\tau} = \frac{\int_t^{t+\tau} f(0, s)ds}{\tau} + \sigma\frac{1 - e^{-\frac{\kappa}{2}\tau}}{\frac{\kappa}{2}\tau}x(t) - \frac{v_1(t, \tau)}{\tau} \tag{5.63}$$

To express this term structure in terms of the value of the instantaneous short rate, expression (5.56) is used to obtain

$$R(t,\tau) = -\frac{\ln P(t,\tau)}{\tau} = \frac{\int_t^{t+\tau} f(0,s)ds}{\tau} + \frac{1-e^{-\frac{\kappa}{2}\tau}}{\frac{\kappa}{2}\tau} \times \tag{5.64}$$

$$(r(t) - f(0,t)) + v_2(t,\tau)$$

with

$$v_2(t,\tau) = \sigma^2 \frac{1-e^{-\frac{\kappa}{2}\tau}}{\frac{\kappa}{2}\tau} \frac{2}{\kappa^2}(1-e^{-\frac{\kappa}{2}t})^2 - \frac{v_1(t,\tau)}{\tau} \tag{5.65}$$

The value of a particular yield at time t [0, T] with maturity τ is the sum of the difference between the short rate and the instantaneous spot rate multiplied by a factor reflecting the mean reversion, the forward rate at the initial trade covering the period [t, $t+\tau$] and a term reflecting the mean reversion and the convexity between interest rates and bond prices. At first sight, this expression does not appear to have any relation to the Vasicek model, which is supposed to be the analogue. However, if, for example, the initial yield curve is flat, the term structure of interest rates can be simplified to

$$R(t,\tau) = R(t,\infty) + \frac{1-e^{-\frac{\kappa}{2}\tau}}{\frac{\kappa}{2}\tau}(r(t) - R(t,\infty)) + v_2(t,\tau) \tag{5.66}$$

As in the original Vasicek model, yields are a weighted combination of the instantaneous spot rate and the infinite maturity yield, with positive weights summing to one plus some function. This latter function is the product of interest rate volatility and another weighting function, which is zero at both ends of the maturity spectrum, causing a curvature of the yield curve. Although the initial yield curve is flat, future term structures of interest rates are curved.

The value of a European call option $C(K, \tau_1, \tau_2)$ at the initial valuation date with maturity τ_1 written on a discount bond with maturity τ_2 is now,

$$C(K,\tau_1,\tau_2) = E_Q\left(e^{-\int_0^{\tau_1} r(s)ds} Max(P(\tau_1, \tau_2 - \tau_1) - K, 0)\right) \tag{5.67}$$

yielding

$$C(K,\tau_1,\tau_2) = P(0,\tau_2)N(d_1) - KP(0,\tau_1)N(d_2) \tag{5.68}$$

with

$$d_1 = \frac{1}{v}\ln\left(\frac{P(0, \tau_2)}{KP(0, \tau_1)}\right) + \frac{1}{2}v$$

$$d_2 = d_1 - v$$

$$v^2 = \sigma^2 \frac{4}{\kappa^3}((1 - e^{-\frac{\kappa}{2}(\tau_2 - \tau_1)})^2 - (e^{-\frac{\kappa}{2}\tau_2} - e^{-\frac{\kappa}{2}\tau_1})^2)$$

Although the initial bond prices may differ, this option pricing formula is equal to the one derived in Section 5.1 where the model of Vasicek was examined.

In the specification of the continuous time economy, investors agree on the unique probability measure Q. The term structure of interest rates and the values of interest rate derivative securities are obtained by taking the appropriate expectations, thus ensuring that relative values of securities are martingales and excluding riskless arbitrage opportunities. As noted before, investors do not have to agree on the actual probability measure in the continuous time economy. To investigate the actual possible distributions of future interest rates resulting from different equivalent probability measures, the Radon-Nikodym derivative $\rho=\{\rho(\lambda, t), t \quad [0, T]\}$ is again defined by

$$\rho(\lambda, t) = exp\left(\int_0^t \lambda d\tilde{W}(s) - \frac{1}{2}\int_0^t \lambda^2 ds\right) \tag{5.69}$$

For ease of exposition, the market price of risk, λ, is assumed to be a fixed-real valued constant.[17] By Girsanov's Theorem, the stochastic process $W=\{W(t), t \quad [0, T]\}$, defined by

$$W(t) = \tilde{W}(t) - \int_0^t \lambda ds \tag{5.70}$$

is a Standard Brownian Motion on the probability space $(\Omega, \mathcal{F}, F, P)$, where the equivalent probability measure P is given by

$$dP = \rho(\lambda, t)dQ \tag{5.71}$$

Given some fixed value of λ, interest rates at time $t \quad [0, T]$, given the information at the initial trade date, are normally distributed with respect to the probability measure P

$$R(t, \tau) \mid \mathcal{F}_0 \sim N(\mu_R(t), \sigma_R^2(t)) \tag{5.72}$$

with

$$\mu_R(t) = \frac{\int_t^{t+\tau} f(0, s)ds}{\tau} + \sigma\lambda t \frac{1 - e^{-\frac{\kappa}{2}\tau}}{\frac{\kappa}{2}\tau} - \frac{v_1(t, \tau)}{\tau} \tag{5.73}$$

and

$$\sigma_R^2(t) = \sigma^2 \left(\frac{1 - e^{-\frac{\kappa}{2}\tau}}{\frac{\kappa}{2}\tau}\right)^2 \left(\frac{1 - e^{-\kappa t}}{\kappa}\right) \tag{5.74}$$

In contrast to the Ho and Lee model, the mean of future interest rates converges to the infinite maturity yield at the initial trade date as maturity increases. The term structure of interest rate volatilities is the same as in the Vasicek model and converges to zero.[18]

The Heath, Jarrow and Morton II Model

The previous two sections dealt with the model of Ho and Lee (1986) and a particular version within the class of Heath *et al.* (1992) term structure models. These models are a counterpart of the models of Merton (1973) and Vasicek (1977). Similar to these two models, interest rates are normally distributed and negative values can occur with a strictly positive probability. A natural step in our presentation of the exogenous term structure of interest rates models is an illustration of the analogue of the CIR (1985) model. Although this is extensively discussed in Heath *et al.* (1992, Section 8), further attention will not be given to it in this section for obvious reasons.

The CIR model incorporates some reasonable features of the possible stochastic movement of the short rate like mean reversion and the exclusion of negative values. In addition, the endogenous term structure of interest rates and the time series characteristics allow for closed-form solutions and as a result, it is very popular from an empirical point of view. The exogenous counterpart of this model, however, does not result in closed-form solutions for interest rate distributions, and the exclusion of negative interest rates restricts the possible shape of the initial term structure of interest rates. Because there are alternative ways of modelling the volatility of the instantaneous short rate to exclude negative interest rates, one of which is discussed in detail analytically and empirically in Heath *et al.* (1990b, 1992), discussion of this other model will take place here.

Because the class of interest rates models of Heath *et al.* (1992) is very general and includes all exogenous term structure of interest rates approaches to be discussed in this thesis, the illustration of this model is also more general than in the previous two sections.

The continuous time uncertainty of the risk-neutral economy is again specified by the filtered probability space $(\Omega, \mathcal{F}, F, Q)$, satisfying the usual conditions. There is a trading interval [0, T] with the final trading date $T > 0$ fixed. The term structure of forward rates $f(0, t)$, t [0, T] at the initial trade date 0 is a known, non-random, measurable and absolute integrable function.

The forward rate at time t maturing at time s, $t < s$, t, s [0, T], which is defined as

$$f(t,s)=\frac{\delta \ln P(t,s-t)}{\delta s} \tag{5.75}$$

obeys the following stochastic differential equation:

$$f(t,s)=f(0,s)+\int_0^t \alpha(v,s,f(v,s))dv+\int_0^t \sigma(v,s,f(v,s))d\tilde{W}(v) \tag{5.76}$$

The function $\alpha(.,.,.)$ and $\sigma(.,.,)$ are Lipschitz continuous, bounded and satisfy

$$\int_0^T |\,\alpha(t,s,f(t,s))\,|\,dt<\infty \tag{5.77}$$

$$\int_0^T \sigma(t,s,f(t,s))^2 dt<\infty \tag{5.78}$$

with probability one for fixed, but arbitrary s [0, T]. At this point, this volatility function is left unspecified to preserve the general treatment of the exogenous term structure of interest rates models. As mentioned in the introduction to this chapter, a specific volatility function ensuring positive interest rates will be discussed later on. Finally, the process $\tilde{W}=\{\tilde{W}(t),\, t\in[0,T]\}$ is a familiar Standard Brownian Motion initialized at zero.

In the discussions of interest rates models thus far, only the stochastic dynamics of the instantaneous short rate were taken into consideration. It is important to realize that the major results and the derivation of general interest rate model of Heath *et al.* also depend on the characteristics of the short rate only. As will be seen in the remainder of this section, the presentation is considerably facilitated by modelling the entire forward rate curve instead of the short rate only.

The next step is the derivation of the unspecified function $\alpha(.,.)$. The stochastic differential equation of the forward rate curve has to ensure that at the initial trade date, values of discount bonds determined by means of discounted expectation equal their given values. In addition, relative values of interest rate derivative securities have to be martingales with respect to the unique probability measure Q. The value of a discount bond $P^*(t, s-t)$ at some time t [0, T] with maturity $s-t$ is defined by

$$P(t,s-t)=e^{-\int_t^s f(t,v)dv} \tag{5.79}$$

and the value of the money market account $B(t)$, t [0, T] is given by

$$B(t)=e^{\int_0^t r(v)dv} \tag{5.80}$$

Refer to Heath *et al.* (1992, Conditions C.2–C.3) for regularity conditions regarding the stochastic processes of bond prices and money market account.

The stochastic differential equation of which the relative bond price $P^*(t, s-t)=P(t, s-t)/B(t)$ is a strong solution, is now

$$\frac{dZ^*(t,s-t)}{Z^*(t,s-t)} = -\int_t^s \alpha(t,v)dvdt + \frac{1}{2}\left(\int_t^s \sigma(t,v)dv\right)^2 dt - \int_t^s \sigma(t,v)dvd\tilde{W}(t) \tag{5.81}$$

where the dependence of the functions $\alpha(.,.,.)$ and $\sigma(.,.,.)$ on forward rates is suppressed for notational convenience. It is easy to verify that, at the initial trade date, bond prices obeying these specifications are consistent with the given term structure of interest rates or forward rates.

To exclude arbitrage opportunities, the process $Z^*(t, s)$ has to be a martingale with respect to the unique probability measure Q. To accomplishthis, take the function $\alpha(t, s)$, $t < s, t, s$ [0, T] as the solution to

$$\int_t^s \alpha(t,v)dv = \frac{1}{2}\left(\int_t^s \sigma(t,v)dv\right)^2 \tag{5.82}$$

yielding, by taking the derivative with respect to s,

$$\alpha(t,s) = \sigma(t,s)\int_t^s \sigma(t,v)dv \tag{5.83}$$

From the regularity conditions regarding the volatility function $\sigma(t, s), t < s, t, s$ [0, T], it can be concluded that relative bond prices are martingales.[19] It is easy to verify that the models discussed in the previous two sections are particular examples of this general class of interest rates models.

As noted earlier, the general derivation of conditions excluding arbitrage opportunities is facilitated considerably by modelling the entire term structure of forward rates. At every time t [0, T], the value of all forward rates is known if the particular path of the Standard Brownian Motion up to this time is known also. The term structure of interest at this time is then by definition

$$R(t,\tau) = \frac{\int_t^{t+\tau} f(t,v)dv}{\tau} = \frac{\int_t^{t+\tau} f(0,v)dv}{\tau} + \frac{\int_t^{t+\tau}\int_0^t \sigma(v,s)\int_t^s \sigma(v,y)dydvds}{\tau} + \frac{\int_t^{t+\tau}\int_0^t \sigma(v,s)d\tilde{W}(v)ds}{\tau} \tag{5.84}$$

The value of a European call option $C(K, \tau_1, \tau_2)$ at the initial trade date with maturity τ_1 written on a discount bond with maturity τ_2 is derived by taking the appropriate expectations, that is,

$$C(K, \tau_1, \tau_2) = E_Q\left(e^{-\int_0^{\tau_1} r(s)ds} Max(P(\tau_1, \tau_2 - \tau_1) - K, 0)\right) \tag{5.85}$$

Because no elegant presentation of this expectation is possible, the next chapter will discuss numerical methods to value interest rate derivative securities.

The presentation of the general arbitrage-free valuation of interest rate derivative securities when the initial term structure of interest rates is exogenously specified, has been similar to the previous sections. Investors agree on the unique probability measure Q, under which relative values of all traded securities are martingales. As a result, arbitrage opportunities are excluded and investors are risk-neutral in this continuous time economy. In the actual economy, however, investors do not have to agree on the probability measure; because of this the possible distributions of future term structures of forward rates will next be examined. First, the familiar Radon-Nikodym derivative $\rho=\{\rho(\lambda, t), t \quad [0, T]\}$ will be defined as

$$\rho(\lambda, t) = exp\left(\int_0^t \lambda d\tilde{W}(s) - \frac{1}{2}\int_0^t \lambda^2 ds\right) \tag{5.86}$$

Assume that the market price of risk, λ, is a real-valued constant.[20] By Girsanov's Theorem, the stochastic process $W=(W(t), t \quad [0, T]\}$, defined by

$$W(t) = \tilde{W}(t) - \int_0^t \lambda ds \tag{5.87}$$

is a Standard Brownian Motion on the probability space $(\Omega, \mathcal{F}, F, P)$, where the equivalent probability measure P is given by

$$dP = \rho(\lambda, t)dQ \tag{5.88}$$

For some fixed value of λ, the forward rate curve at some time $t \quad [0, T]$, given some realization of the Standard Brownian Motion $W=\{W(t), t \quad [0, T]\}$ up to this time, is

$$f(t, s) = f(0, s) + \int_0^t \sigma(v, s)\int_v^s \sigma(v, y)dydv + \lambda\int_0^t \sigma(v, s)dv + \tag{5.89}$$

$$\int_0^t \sigma(v, s)dW(v)$$

Until now, the volatility function has not been specified and, as noted before, all interest rates models discussed are a special case within this general class of interest rates models. To ensure that negative values of future interest rates are excluded, take, as in Heath *et al.* (1992, p. 95),

$$\sigma(t, s, f(t, s)) = \sigma Min(f_{Max}, f(t, s)) \tag{5.90}$$

with σ and f_{Max} positive, real-valued constants. The volatility is proportional to the value of the forward rate, when forward rates are “small”, and is constant, when forward rates are “large”. As shown in Heath *et al.* (1992, Proposition 5), this specific volatility function guarantees the existence of a solution to the stochastic differential equation (5.76). In addition, negative interest rates are excluded with probability one.[21,22]

Exogenous Term Structure of Volatility Models

The main difference between the endogenous and exogenous term structure of interest rates models is the specification of the parameters of the stochastic process of the instantaneous spot rate. Within the class of endogenous term structure models, the parameters are assumed to be constant. The exclusion of arbitrage opportunities then implies a specific yield curve at some valuation date. The models of Merton (1973) Vasicek (1977) and Cox *et al.* (1985) are some well-known examples of this approach and are discussed in previous sections.

Because the drift term of the stochastic process is time-dependent, the term structure of interest rates at the valuation can be any arbitrary function. The volatility of the short rate implies only a term structure of interest rate volatilities, whereas in the above-mentioned approach, both term structures were dependent on this parameter. In the previous sections, the model of Ho and Lee (1986) and two versions of the general class of interest rates models of Heath *et al.* (1992) were discussed. The first two of these models are counterparts of the Merton and Vasicek model, while the last one is comparable to the model of Cox *et al.*, as negative interest rates are excluded.

The main similarity between the models discussed so far is the implied term structure of interest rate volatilities. Because the stochastic dynamics of the instantaneous short rate are driven by a Standard Brownian Motion and a few constant parameters, this term structure at the valuation date is also an implied function of these parameters. It is not surprising, therefore, that by allowing the short rate volatility to be dependent on calendar time, the term structure of interest rate volatilities can be an arbitrary or, in other words, a given function. In Chapter 3, it was already noted that this extension or generalization does not mean that this class of models, in which both term structures are exogenous, is able to describe the actual stochastic dynamic behavior of interest rates over time more accurately. In forthcoming chapters, where some empirical results regarding term structures of interest rates and the valuation of derivative securities will be extensively discussed, it is argued that the large amount of freedom created by this time-dependency of the coefficients may actually lead to estimation problems.

In principle, the analysis of the different interest rates models of the previous sections can be repeated for the case of a time-varying volatility. In this section, however, only the extended version of the Ho and Lee model, derived by Jamshidian (1990)[23] will be discussed. Because this model is empirically analyzed, a detailed treatment of the various stochastic characteristics and valuation formulas seems to be justified. In addition, the extended square-root process of Hull and White (1990b) or the extended log-normal process of Black *et al.* (1990) generally leads to too many parameters or unknown distributions of interest rates and corresponding estimation problems. To maintain the general treatment of interest rates models, refer to the discussion of the general class of models of Heath *et al.* (1992).

The Jamshidian Model

As in the analysis of the original Ho and Lee model, the continuous time uncertainty of the risk-neutral economy is specified by the filtered probability space $(\Omega, \mathcal{F}, F, P)$, satisfying the usual conditions, and a trading interval [0, T] with $T > 0$ fixed. Investors agree on the unique probability measure Q. The exogenous term structure of forward rates at the initial trade date is given by $f(0, t)$, t [0, T], which is a non-random, measurable and absolute integrable function. At this point, the exogenous term structure of interest rate volatilities is not specified. Because the general shape of this term structure is derived from the model specification, a discussion of some limitations can take place; an explicit characterization is thus premature.

The value of the instantaneous short rate at time t [0, T] obeys

$$r(t) = \alpha(t) + \sigma(t)\tilde{W}(t) \tag{5.91}$$

The functions $\alpha(t)$ and $\sigma(t)$ are deterministic, bounded and Lipschitz continuous. The process $\tilde{W} = \{\tilde{W}(t), t \in [0, T]\}$ is again a Standard Brownian Motion initialized at zero.

The first step in the discussion is the characterization of $\alpha(t)$, t [0, T]. Similar to previous presentations, the given term structure of interest rates at the initial trade date 0 and the explicit requirement of relative values of securities being a martingale under the unique probability measure Q, results in an analytic solution to the unknown function $\alpha(t)$, t [0, T]. The value of a discount $P(0, \tau)$ at time 0 with maturity τ, τ [0, T], therefore, necessarily equals the following two expressions

$$P(0, \tau) = e^{-\int_0^\tau f(0,s)ds} = E_Q\left(e^{-\int_0^\tau r(s)ds} \mid \mathcal{F}_0\right) \tag{5.92}$$

Calculation of the expectation results in[24]

$$\alpha(t) = f(0, t) + \sigma(t)\int_0^t \int_u^t \sigma(s)dsdu \tag{5.93}$$

The value of a discount bond $P(t, \tau)$ at some future point in time t [0, T], can be obtained analogously, yielding

$$P(t,\tau)=E_Q\left(e^{-\int_t^{t+\tau} r(s)ds} \mid \mathcal{F}_t\right)= \tag{5.94}$$

$$exp\left(-\int_t^{t+\tau} f(0,s)ds-\frac{\int_t^{t+\tau}\sigma(s)ds}{\sigma(t)}(r(t)-f(0,t))-\right.$$

$$\left.\frac{1}{2}t\left(\int_t^{t+\tau}\sigma(s)ds\right)^2\right)$$

The term structure of interest rates at this time is, therefore,

$$R(t,\tau)=\frac{\int_t^{t+\tau} f(0,s)ds}{\tau}+\frac{\int_t^{t+\tau}\sigma(s)ds}{\tau}\,\frac{(r(t)-f(0,t))}{\sigma(t)}+ \tag{5.95}$$

$$\frac{1}{2}t\frac{\left(\int_t^{t+\tau}\sigma(s)ds\right)^2}{\tau}$$

Opposed to the nested Ho and Lee model, changes in the short rate do not necessarily imply parallel shifts in the yield curve. In fact, a whole family of possible movements can be obtained after a suitable specification of the volatility function *σ(t), t* [0, *T*]. If forward rates are bounded, the infinity maturity yield is a constant if and only if

$$\lim_{t\to\infty}\sigma(t)=0 \tag{5.96}$$

Although the continuous time economy is not originally designed to investigate infinite maturity yields, it is clear that reasonable values of this yield to remove a basic drawback of the Ho and Lee model can only be obtained by an unreasonable long-term stochastic behavior of the short-term rate of interest.[25]

The value of a call option $C(K, \tau_1, \tau_2)$ at the initial valuation date with maturity τ_1 and underlying discount bond maturing at τ_2 is

$$C(K,\tau_1,\tau_2)=E_Q\left(e^{-\int_0^{\tau_1} r(s)ds}Max(P(\tau_1,\tau_2-\tau_1)-K,0)\right)= \tag{5.97}$$

$$P(0,\tau_2)N(d_1)-KP(0,\tau_1)N(d_2)$$

with

$$d_1 = \frac{1}{v}\ln\left(\frac{P(0,\tau_2)}{KP(0,\tau_1)}\right) + \frac{1}{2}v$$

$$d_2 = d_1 - v$$

$$v^2 = \tau_1\left(\int_{\tau_1}^{\tau_2}\sigma(s)ds\right)^2$$

In order to illustrate the rich pattern of possibilities, option values based on two different volatility functions will be compared to those in the original Ho and Lee model. Simply assume that the implied term structure of interest rates of the Merton model is the same as the exogenous yield curve at the valuation date in the Ho and Lee model. As the volatility parameter of the option formula is the same as in the Merton model, refer to Figure 5.3 for the different base-case scenarios and option and bond characteristics.

The first volatility function $\sigma(t)$, t [0, T] is parabolic in calendar time. At the minimum and maximum maturity of the option, that is, 0 and 10 years, the volatility of the short rate equals 0.04, the value of the base-case scenario. When the time-to-maturity of the option is 5 years, the volatility function attains its maximum value, which is 0.07. The second volatility function is also parabolic in calendar time. However, when the option maturity is 5 years, volatility attains a minimum value of 0.01. It is clear that these volatility functions are not a realistic description of the actual volatility of the short rate. As mentioned before, they only serve as a nice illustration. Figure 5.10 presents the option values.

When we compare this figure to Figure 5.3, it is seen that option values based on the high and low scenarios are generally lower and higher, respectively, than the corresponding scenarios in the case of the Ho and Lee model. What is interesting to examine in this case is the particular shape of the option value as a function of maturity. In the low scenario, option values are rather constant, as maturity varies between two and seven years. In the case of the high scenario, option values are changing heavily within this region. Although both shapes can be explained by the typical functional form of volatility, it is still interesting to see what different kinds of option values can be modelled.

Similar to the discussion of the previous term structure of interest rates models, option valuation formulas and possible future term structures are derived in an arbitrage-free economy. Investors agree on the unique probability measure Q, and relative values of all securities are martingales with respect to this measure. We now derive the distribution of interest rates in the case of different equivalent probability measures, reflecting the fact that investors do not have to agree on actual assessments of certain events. The Radon-Nikodym derivative $\rho=\{\rho(\lambda, t), t$ [0, T]$\}$ is, therefore, defined as

$$\rho(\lambda, t) = exp\left(\int_0^t \lambda d\tilde{W}(s) - \frac{1}{2}\int_0^t \lambda^2 ds\right) \tag{5.98}$$

For simplicity, assume that the market price of risk, λ, is a real-valued constant.[26] By Girsanov's Theorem, the stochastic process $W=\{W(t), t \quad [0, T]\}$, defined by

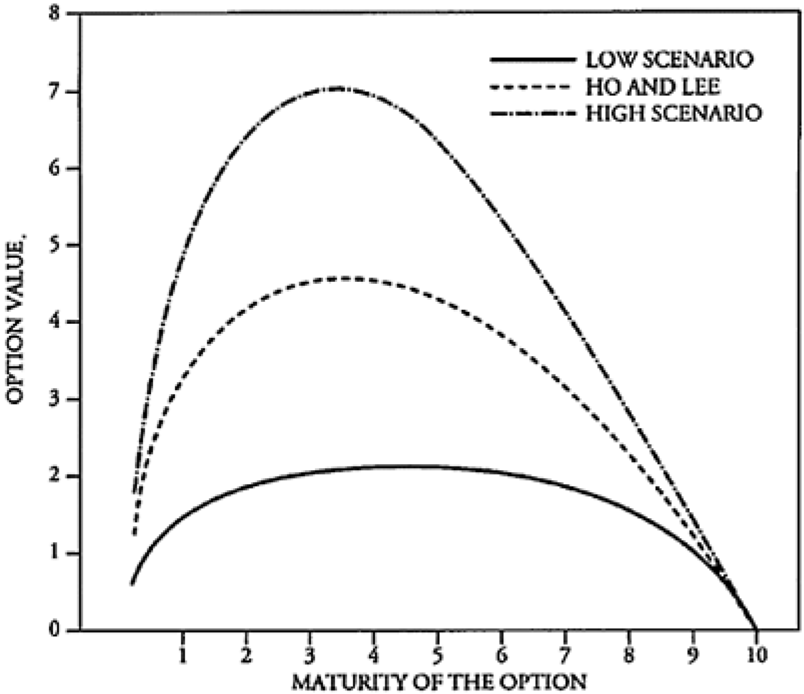

Figure 5.10 European call options on discount bonds

This figure shows the value of a European call option on a discount bond with face value 100 for different volatility functions and for different maturities of the option under the Jamshidian model. The maturity of the underlying bond is ten years and the exercise price of the options is equal to the forward price of the underlying bond. The exogenous term structure of interest rates at the valuation date is equal to the yield curve implied by the Merton model. In this case, the instantaneous short-rate is 0.07, while the risk-neutral drift term and constant volatility equal 0.02, and 0.04, respectively. The volatility functions in the case of the high and low scenario are parabolic functions with value 0.04 at both ends of the maturity spectrum, and with maximum and minimum value of 0.07 and 0.01, respectively.

$$W(t) = \tilde{W}(t) - \int_0^t \lambda ds \qquad (5.99)$$

is a Standard Brownian Motion on the probability space $(\Omega, \mathcal{F}, F, P)$, with P given by

$$dP = \rho(\lambda, t)dQ \qquad (5.100)$$

Given this fixed λ, interest rates at time t, t [0, T], are normally distributed

$$R(t, \tau) \mid \mathcal{F}_0 \sim N(\mu_R(t), \sigma_R^2(t)) \qquad (5.101)$$

with

$$\mu_R(t) = \frac{\int_t^{t+\tau} f(0,s)ds}{\tau} + \frac{\int_t^{t+\tau} \sigma(s)ds}{\tau} \lambda t + \qquad (5.102)$$

$$\frac{\int_t^{t+\tau} \sigma(s)ds}{\tau} \sigma(t) \int_0^t \int_u^t \sigma(s)dsdu + \frac{1}{2}\frac{t}{\tau}\left(\int_t^{t+\tau} \sigma(s)ds\right)^2$$

and

$$\sigma_R^2(t) = \frac{\left(\int_t^{t+\tau} \sigma(s)ds\right)^2 t}{\tau^2} \qquad (5.103)$$

The term structure of interest rate volatilities can attain a number of different shapes. Given such an exogenous term structure, one can solve for the volatility function $\sigma(t)$, t [0, T] and value interest rate derivative securities. It is obvious, however, that the exogenous term structure of interest rate volatilities has to be positive for all maturities or negative for all maturities.[27,28]

NOTES

1 It is easy to see that a multi-factor extension of this model does not make any sense. Suppose there are n factors $x_i(t)$, i=1, ..., n, obeying

$$dx_i(t) = \theta_i dt + \sigma_i dW_i(t)$$

with θ_i and σ_i real-valued constants and $W = \{W_1(t),..., W_n\}$ a vector process of n independent Standard Brownian Motions. If the instantaneous interest rate $r(t)$ is equal to the sum of these n factors, $r(t)$ is again normally distributed, given the information at time s, $s \leq t$, with mean $r(s) + \theta(t - s)$ and variance $\sigma^2(t - s)$ where

$$\theta = \sum_{i=1}^{i=n} \theta_i \quad \text{and} \quad \sigma^2 = \sum_{i=1}^{i=n} \sigma_i^2$$

From this, it can be concluded that the multi-factor extension simplifies to the one-factor case.

2 For a detailed description of the numerical methods used to calculate these option values, see Chapter 6.

3 To review briefly the multi-factor extension, suppose there are n factors $x_i(t)$, $i = 1, \ldots, n$, obeying

$$dx_i(t) = \kappa_i(\theta_i - x_i(t))dt + \sigma_i dW_i(t)$$

with K_i, θ_i, and σ_i real-valued constants and $W=\{W_1(t),\ldots, W_n(t)\}$ a vector process of independent Standard Brownian Motions. If the instantaneous interest rate $r(t)$ is the sum of these n factors, the analysis in this section can be easily generalized to obtain the value of a discount bond $P_n(t, \tau)$ at time t with maturity τ and unit face value

$$P_n(t, \tau) = \prod_{i=1}^{i=n} P(x_i(t), t, \tau)$$

The derivation of the value of a European call option $C_n(t, K, \tau_1, \tau_2)$ on a discount bond in this case is also similar:

$$C_n(t, K, \tau_1, \tau_2) = P_n(t, \tau_2)N(d_1) - KP_n(t, \tau_1)N(d_2)$$

where

$$d_1 = \frac{1}{v}\ln\left(\frac{P_n(t,\tau_2)}{KP_n(t,\tau_1)}\right) + \frac{1}{2}v$$

$$d_2 = d_1 - v$$

$$v^2 = \sum_{i=1}^{i=n} \frac{1}{2}\frac{\sigma_i^2}{\kappa_i^3}((1 - e^{-\kappa_i(\tau_2-\tau_1)})^2 - (e^{-\kappa_i(\tau_2-t)} - e^{-\kappa_i(\tau_1-t)})^2)$$

For a thorough discussion of the multi-factor extension of the Vasicek (1977) model, see Langetieg (1980).

4 For a thorough discussion of the square root process, see Feller (1951).

5 To review the multi-factor extension of this model briefly, suppose there are n factors $x_i(t)$, i=1,..., n, obeying

$$dx_i(t) = \kappa_i(\theta_i - x_i(t))dt + \sigma_i\sqrt{x_i(t)}dW_i(t)$$

with K_i, θ_i and σ_i real-valued constants and $W=\{W_1(t),..., W_n(t), t \geq 0\}$ a vector process of n independent Standard Brownian Motions. If the instantaneous interest rate $r(t)$ is equal to the sum of these n factors, each of which is ensured to be positive by imposing the familiar parameter constraint, the value of a discount bond $P_n(t, \tau)$ at time t with maturity τ and unit face value is simply

$$P_n(t,\tau) = \prod_{i=1}^{i=n} P(x_i(t), t, \tau)$$

The value of the European call option on a discount bond can be derived similarly.

For a discussion of the two-factor extension of this model, see Cox *et al.* (1985, p. 398–401) and Richard (1978).

6 In Courtadon (1982, p. 92) it is shown that if interest rates follow a geometric Brownian Motion with zero drift,

$$\lim_{t\to\infty} r(t) = 0 \quad \text{with probability one}$$

by the Law of Iterated Logarithms. As a result, it is concluded that this process is inadequate to represent the long-term behavior of interest rates, although it might be a good appoximation in the short run.

7 For a detailed description of the numerical methods used to calculate the values of interest rate derivative securities, see Chapter 6.

8 See Courtadon (1982, p. 98–9).

9 The value of the consol bond $P_l(c, l(t))$ at time t paying a continuous coupon c is

$$P_l(c, l(t)) = \lim_{\tau \to \infty} \int_t^{t+\tau} c e^{-l(t)s} ds = \lim_{\tau \to \infty} \left[-\frac{c}{l(t)} e^{-l(t)s} \right]_t^{t+\tau} = \frac{c}{l(t)}$$

10 Given these specifications, equation (5.43) can also be rewritten as the following stochastic differential equation

$$dr(t) = \frac{\partial \alpha(t)}{\partial t} dt + \sigma d\tilde{W}(t)$$

which is a more familiar representation in comparison to the previous sections.

11 To evaluate this expression, we have used

$$E(e^{\tilde{x}}) = e^{E(\tilde{x}) + \frac{1}{2} Var(\tilde{x})}$$

if $\tilde{x}$ is normally distributed.

12 Another way of explaining this phenomenom is to compare the initial term structures of forward rates. In the Merton model, forward rates are decreasing without bound as maturity increases.

13 For a detailed exposition of these necessary requirements, see Heath *et al.* (1992, Condition C.4).

14 Similar to the Merton (1973) model, a multi-factor extension does not make any sense.

15 This derivation is similar to de Munnik (1992) and facilitates the numerical analysis; see Chapter 6.

16 See n. 11.

17 See again Heath *et al.* (1992, Condition C.4).

18 As is the case with the Vasicek (1977) model, a multi-factor extension is easily accomplished. Suppose there are n factors $x_i(t)$, i=1,..., n, obeying

$$x_i(t) = \int_0^t e^{-\frac{\kappa_i}{2}(t-s)} d\tilde{W}_i(s)$$

with K_i real-valued constants and $\tilde{W} = \{\tilde{W}_1(t), \ldots, \tilde{W}_n(t)\}$ a vector process of independent Standard Brownian Motions. A multi-factor extension of the analysis then yields

$$r(t) = f(0, t) + \sum_{i=1}^{i=n} \sigma_i^2 \frac{2}{\kappa_i^2} (1 - e^{-\frac{\kappa_i}{2} t})^2 + \sum_{i=1}^{i=n} \sigma_i x_i(t)$$

where σ_i are real-valued constants. The term structure at some future date t [0, T] is

$$R(t,\tau)=\frac{\int_t^{t+\tau} f(0,s)ds}{\tau}+\sum_{i=1}^{i=n}\sigma_i\frac{1-e^{-\frac{\kappa_i}{2}\tau}}{\frac{\kappa_i}{2}\tau}x_i(t)-\frac{v_{1i}(t,\tau)}{\tau}$$

with $v_{1i}(t, \tau)$ defined accordingly. The value of the option $C_n(K, \tau_1, \tau_2)$ is

$$C_n(K,\tau_1,\tau_2)=P(0,\tau_2)N(d_1)-KP(0,\tau_1)N(d_2)$$

with d_1 and d_2 defined as before and

$$v^2=\sum_{i=1}^{i=n}\sigma_i^2\frac{4}{\kappa_i^3}((1-e^{-\frac{\kappa_i}{2}(\tau_2-\tau_1)})^2-(e^{-\frac{\kappa_i}{2}\tau_2}-e^{-\frac{\kappa_i}{2}\tau_1})^2)$$

19 See Heath *et al.* (1992, Condition C.3) and Duffie(1988, Page 140 and Section 15.E).

20 See Heath *et al.* (1992, Condition C.4) for general regularity conditions regarding the market price of risk.

21 It is important that the volatility of forward rates is constant when forward rates are "large". Suppose

$$\sigma(t,s,f(t,s))=\sigma f(t,s)$$

with σ a positive, real-valued constant. In Morton (1988) it is shown that in this case no finite valued solution to (5.76) exists, because volatility is unbounded.

22 In case of a multi-factor extension of this model, the forward rate curve at some time t [0, T], given n volatility functions $\sigma_i(t, s)$, $t < s$, t, s [0, T], n fixed real-valued market prices of risk ($\lambda_1, \ldots, \lambda_n$) and some realization of the vector process $W=\{W_1(t), \ldots, W_n(t), t$ [0, T]$\}$, is

$$f(t,s)=f(0,s)+\sum_{i=1}^{i=n}\int_0^t\sigma_i(v,s)\int_v^s\sigma_i(v,y)dydv+$$

$$\sum_{i=1}^{i=n}\lambda_i\int_0^t\sigma_i(v,s)dv+\sum_{i=1}^{i=n}\int_0^t\sigma_i(v,s)dW_i(v)$$

As before, see Heath *et al.* (1992) for a detailed overview of the regularity conditions concerning the volatility functions and market prices of risk. Finally, the valuation of interest rate derivative securities is accomplished similarly, by generalizing the one-factor risk-neutral expectation.

23 The binomial version or discretizaton of this model was derived independently by Pederson *et al.* (1989).

24 See n. 11 and Arnold (1974, ch. 8)

25 As in the previous sections, the introduction of a mean reverting process for the short rate also removes this drawback, and preserves reasonable stochastic dynamics. However, the estimation problems already noted in the introduction to this section remain. For a detailed treatment of this process, see Jamshidian (1990).

26 See again Heath *et al.* (1992, Condition C.4) for general regularity conditions regarding the market price of risk.

27 This can be explained by the symmetry of the Brownian Motion.

28 In the case of *n* factors, the short rate obeys

$$r(t) = f(0, t) + \sum_{i=1}^{i=n} x_i(t)$$

with

$$x_i(t) = \sigma_i(t) \int_0^t \int_u^t \sigma_i(s) ds du + \sigma_i(t) \tilde{W}_i(t)$$

and the volatility functions $\sigma_i(t),\ t \in [0, T],\ i = 1, \ldots, n$ satisfying the requirements described above and $W = \{W_1(t), \ldots, W_n(t), t \in [0, T]\}$ a vector process of independent Standard Brownian Motions. The term structure of interest rates at time t [0, T] is

$$R(t, \tau) = \frac{\int_t^{t+\tau} f(0, s) ds}{\tau} + \sum_{i=1}^{i=n} \frac{\int_t^{t+\tau} \sigma_i(s) ds}{\tau} \frac{x_i(t)}{\sigma_i(t)} + \frac{1}{2} t \sum_{i=1}^{i=n} \frac{\left(\int_t^{t+\tau} \sigma_i(s) ds\right)^2}{\tau}$$

The value of the option $C_n(K, \tau_1, \tau_2)$ is

$$C_n(K, \tau_1, \tau_2) = P(0, \tau_2) N(d_1) - K P(0, \tau_1) N(d_2)$$

with d_1 and d_2 defined as before and

$$v^2 = \tau_1 \sum_{i=1}^{i=n} \left(\int_{\tau_1}^{\tau_2} \sigma_i(s) ds \right)^2$$

For a detailed exposition and empirical application of this model, see Jamshidian (1989) and Chapter 9.

6
NUMERICAL METHODS TO VALUE INTEREST RATE DERIVATIVE SECURITIES

Chapter 2 discussed the general issue of the valuation of interest rate derivative securities. The most important result of this chapter is the equivalence between the possibility of valuing any security or claim uniquely and arbitrage-free and the existence of a unique probability measure such that prices expressed in terms of a money-market account are martingales. If this condition is fulfilled, the value of such a security has to be equal to the expected discounted value of the claim. The expectation in this case should be taken with respect to the unique probability measure.

In previous chapters different approaches for valuing specifically interest rate dependent securities were categorized and a variety of well-known models developed in recent years were discussed. The two main categories are characterized by the explicit formulation of the underlying value of an interest rate contingent claim. In the case where the value of a bond is treated as a state variable and exogenously or directly modelled, the resulting valuation procedure is classified as the direct approach. The indirect approach, however, considers the value of a discount bond as an interest rate contingent claim, too. In the modelling of the instantaneous short-term rate of interest, the value of the bond or the term structure of interest rates is derived by taking the familiar discounted expectation.

Most of the models discussed so far have been illustrated by calculating the value of a European call option written on a discount bond. The reasons for concentrating the analysis on this particular interest rate contingent claim are twofold. Within the field of valuing interest rate derivative securities, this claim has attracted a lot of attention because it has a valuation problem similar to that of options on stocks. As a result, an interesting comparison could be made between these two claims. In addition, for most interest rates models, the value of a European call option written on a discount bond can be expressed as a closed-form solution, facilitating the above-mentioned analysis and clearly illustrating the general arbitrage-free valuation procedure.

A number of interest rate derivative securities mentioned in Chapter 1, however, do not allow for the derivation of a closed-form solution that can readily be calculated. In order to determine the values of these instruments, therefore, numerical procedures have to be used. These procedures have the common property of evaluating sample paths of the underlying stochastic state variable. Based on these particular paths and their corresponding probabilities, the value of the contingent claim is again obtained by calculating the discounted expectation.

A straightforward approach that utilizes this evaluation of sample paths is simulation. In the case of the Black and Scholes (1973) model, for example, the value of a European call option on a non-dividend paying stock can be determined by randomly generating stock prices that are risk-neutrally distributed at maturity of the option, and calculating the discounted average option value. By increasing the number of stock prices generated, a reasonable approximation of the option is obtained.[1] Although this method can handle complicated distributions of the state variables, its inability to incorporate early-exercise features of contingent claims still severely limits the method's practical application.

A lattice approach, however, does incorporate this early-exercise feature because of the ability to represent the conditional distribution of the underlying state variables. Each point of the lattice represents the value of the state variable at some point in time in some state. From this particular point, different nodes of the lattice, representing different states at some future time, can be reached with some probability. The actual values at each node and the corresponding probabilities are chosen such that by increasing the number of steps, the discrete-time distribution converges to the continuous-time distribution. Based on such a tree, the value of the contingent claim is calculated in the usual way.

The lattice approaches can be classified according to the extent to which the different sample paths recombine. If nodes representing possible future values of the state variable can be reached with positive probability by several previous nodes, the lattice, or tree, is path-independent. It is obvious from this definition that if such nodes can only be reached by one previous node, the sample paths of the tree do not recombine and, thus, the lattice is denoted path-dependent.

This categorization of the lattice approaches might suggest that path-independent trees are generally preferable. Given a fixed number of points in time or dates to span some time interval, the corresponding total number of nodes in this approach is less than in the path-dependent approach, facilitating numerical application and possibly increasing convergence and accuracy. As will be seen in the next sections, however, the constraints imposed on the lattice to let the paths recombine can decrease the rate of convergence. In addition, the distributional characteristics of some interest rate processes are dependent on the particular realizations of previous interest rates and, as a result, the construction of path-independent trees might not be possible.

The first section of this chapter discusses three general approaches to constructing path-independent interest rate trees. Although the analysis is primarily concerned with the valuation of interest rate derivative securities, the approaches are easily extendable and applicable for different state variables. Apart from this detailed treatment, a comparison of the different methods will be made in the second section by actually calculating options on bonds for the Cox *et al.* (1985) model. In addition to this practical demonstration, some criteria to assess and compare the rate of convergence and numerical stability will be put forward. After this, the third section analyzes the model of Heath, Jarrow and Morton I and shows the construction of a path-independent interest rate tree.

In general, path-dependent trees are constructed only for specific cases, when another approach is not applicable because of distributional characteristics. A well-known example is the particular version of the model of Heath, Jarrow and Morton II, discussed in Chapter 5. Because of the popularity of this model in financial practice and the fact that in the empirical part of this thesis some serious objections to it are discussed with respect to the empirical implementation, the construction of a path-dependent interest rate tree is discussed in the last section.

PATH-INDEPENDENT NUMERICAL METHODS

If some nodes representing possible future values of the interest rate can be attained by more than one previous node with positive probability, the lattice, or tree, is path-independent. This recombining of sample paths reduces the total number of nodes in the tree compared to the path-dependent approach and there are obvious numerical reasons mentioned in the introduction to this chapter for constructing these modified trees.

The desired reduction of nodes, however, imposes some constraints on the drift and volatility functions of the stochastic interest rate process. In this section, these conditions are formulated and three well-known approaches for constructing a path-independent interest rate tree are discussed. The first numerical procedure to be discussed is the method of Nelson and Ramaswamy (1990). In this approach, a binomial tree is constructed to represent the possible values of future interest rates. By modifying the probabilities at each node of the tree, this approach is able to represent a general class of stochastic processes (to be defined later on). Although the path-independent tree is binomial, multiple jumps may be necessary to accomplish the convergence of the discrete-time distribution to the continuous-time equivalent. In the method of Tian (1991), for instance, these multiple jumps are excluded, simplifying the numerical procedure. Finally, the method of Hull and White (1990a), who construct a trinomial tree based on the explicit finite difference method, will be analyzed.

The first step in the construction of the interest rate lattice in these approaches is the following necessary transformation of the stochastic process of the interest rate. Suppose the general stochastic differential equation of the instantaneous short-term rate of interest in the risk-neutral economy is given by

$$dr(t) = \mu_r(r(t), t)dt + \sigma_r(r(t), t)d\tilde{W}(t) \tag{6.1}$$

The initial value of the interest rate $r(0)$ is a constant and the process $\tilde{W} = \{\tilde{W}(t), t \in [0, T]\}$ is a Standard Brownian Motion initialized at zero. The assumption is made that a solution to this equation exists with probability one, which is distributionally unique. The functions $\mu_r(r(t), t)$ and $\sigma_r(r(t), t)$ are continuous and are dependent only on the information at time t, represented by the value of the interest rate at this time, and not on the particular path followed by the interest rate through $[0, t]$. As will be seen in the remainder of this section, this assumption is crucial in order to obtain a path-independent interest rate tree that will approximate the continuous process. In addition, the drift and volatility functions obey the following conditions:[2]

$$\mu_r(0, t) \geq 0$$

$$\sigma_r(0, t) = 0$$

Although this singularity at $r(t) = 0$, t [0, T], limits, to some extent, our general treatment, it is a reasonable assumption for the modelling of asset prices or nominal interest rates within the model of Cox *et al.* (1985). In addition, the algorithm presented below is easily modified to deal with processes without this singularity.

The basic principle of the numerical tree consists of the discrete approximation of the Standard Brownian Motion and the transformation at each node of the tree to obtain a corresponding interest rate value. After this, suitable probabilities have to be specified to match the instantaneous drift and volatility functions. If, however, this procedure is applied to equation (6.1), the tree is not path-independent if the volatility function is explicitly dependent on the value of the interest rate. To avoid this pathdependency, the original interest rate process must be transformed such that it has an instantaneous standard deviation equal to one. Suppose therefore, that $x(r(t), t)$ is twice differentiable in $r(t)$ and once in t, t [0, T]. Using Ito's Lemma,

$$dx(r(t), t) = \mu_x(x, t)dt + \sigma_x(x, t)d\tilde{W}(t) \tag{6.2}$$

with

$$\mu_x(x, t) = \frac{\partial x}{\partial t} + \frac{\partial x}{\partial r}\mu_r(r, t) + \frac{1}{2}\frac{\partial^2 x}{\partial r^2}\sigma_r^2(r, t)$$

$$\sigma_x(x, t) = \frac{\partial x}{\partial r}\sigma_r(r, t)$$

For notational convenience, the dependency of the variable $r(t)$ on t, t [0, T], has been suppressed. To get a unit-instantaneous standard deviation of this transformed variable $x(r, t)$, t [0, T], take

$$x(r, t) = \int_{r(0)}^{r(t)} \frac{dz}{\sigma_r(z, t)} \tag{6.3}$$

In addition, define

$$x^L(t) = \lim_{r \downarrow 0} x(r, t)$$

$$x^U(t) = \lim_{r \to \infty} x(r, t)$$

and, after assuming both $x^L(t)$ and $x^U(t)$ to be real-valued constants for all t [0, T], the inverse transformation is defined as

$$r(t) = \begin{cases} r : x(r(t), t) = x & \text{if} \quad x^L < x < x^U \\ \infty & \text{if} \quad x^U \leq x \\ 0 & \text{if} \quad x \leq x^L \end{cases}$$

for t [0, T].

The presentation of the different procedures in the following subsections starts with the illustration of the construction of the tree to approximate the transformed process $x=\{x(t), t \in [0, T]\}$. Next, a derivation is made of the exact number of nodes in a tree, given a fixed number of steps to divide the time interval spanned by this tree, as this number is useful in the general comparison in the next section. A demonstration next follows of the different numerical methods, using a simple example of valuing an American call option on a discount bond under the model of Nelson and Ramaswamy (1990).

In Chapter 5, limited attention was paid to the multi-factor extensions of the different interest rates models, as this extension is easily accomplished by making the short-term rate a sum of several independent stochastic variables. For the same reason, the discussion of the numerical procedures in this chapter concentrates on one-factor interest rates models.

The Method of Nelson and Ramaswamy

The stochastic process of the transformed variable $x=\{x(t), t \in [0, T]\}$ is approximated by a binomial tree consisting of n steps. Each of these steps represents a subinterval of length $\Delta t = T/n$, and the different values of the approximate process $\hat{x} = \{\hat{x}(t_j), t_j = j\Delta t, j = 0, \ldots, n\}$ at each point of the lattice are given by

$$\hat{x}_i(t_j) = (j - 2i)\sqrt{\Delta t} \tag{6.4}$$

with $i=0,\ldots, j$ and $j=0,\ldots, n$. In Figure 6.1, the resulting binomial tree is shown; at each point of the tree, the approximate process can either increase or decrease. Use of the inverse transformation yields the corresponding approximate discrete values of the interest rate process $\hat{r} = \{\hat{r}(t_j), t_j = j\Delta t, i = 0, \ldots, n\}$. The binomial probabilities at each point of the lattice are now chosen to match the instantaneous drift term of the continuous-time process given by equation (6.2) exactly. At node (i, t_j) with $i=0,\ldots, j$ and $j=0,\ldots, n-1$, therefore, the probability of an upward move during the next period is[3]

$$q_i(t_j) = \frac{\mu_r(\hat{r}_i(t_j), t_j)\Delta t + \hat{r}_i(t_j) - \hat{r}_{i-1}(t_{j+1})}{\hat{r}_{i+1}(t_{j+1}) - \hat{r}_{i-1}(t_{j+1})} \tag{6.5}$$

The assumed singularity of the instantaneous volatility function in relation to the conditions regarding the instantaneous drift term, however, might result in negative binomial probabilities. By allowing multiple upward and downward jumps in the binomial tree, however, one can avoid this complication. At each point of the lattice, the minimum number of upward jumps required to keep the upward probability less than one is given by

$$I^U(\hat{x}_i(t_j)) = \begin{cases} \text{The smallest, odd, positive integer } k \text{ such that} & \\ \hat{r}_{i+k}(t_{j+1}) - \hat{r}_i(t_j) \geq \mu_r(\hat{r}_i(t_j), t_j)\Delta t & \text{If } \hat{x}_i(t_j) < x^U \\ 1 & \text{If } \hat{x}_i(t_j) \geq x^U \end{cases}$$

Similarly, the minimum number of downward jumps required to keep the upward probability positive is given by

$$I^D(\hat{x}_i(t_j)) = \begin{cases} \text{The smallest, odd, positive integer } k \text{ such that} \\ \text{Either} \quad \hat{r}_{i-k}(t_{j+1}) - \hat{r}_i(t_j) \leq \mu_r(\hat{r}_i(t_j), t_j)\Delta t \\ \text{Or} \quad \hat{r}_{i-k}(t_{j+1}) = 0 \end{cases}$$

This multiple upward and downward jumping to modify the branching process is illustrated in Figure 6.1. At node (−1, t_j), the transformed binomial process can attain either (2, t_{j+1}) or (−2, t_{j+1}). The corresponding values of $I^U(\hat{x}_{-1}(t_i))$ and $I^D(\hat{x}_{-1}(t_i))$ are 3 and −1, respectively.

Based on these possible multiple jumps at each point of the node, the modified binomial upward probability at each node *(i, t_j),* with i=0,…, j and j = 0,…, n−1 is

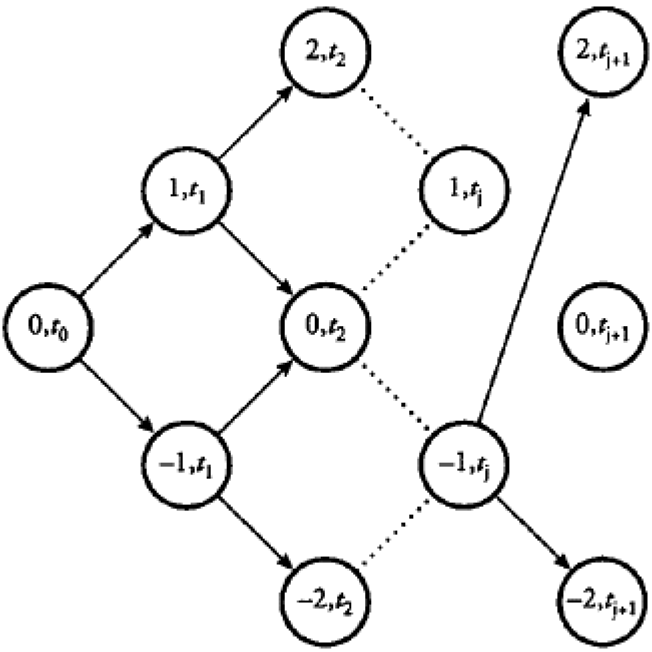

Figure 6.1 The method of Nelson and Ramaswamy

This figure shows the binomial lattice constructed by means of the method of Nelson and Ramaswamy. At each point of the tree, the variable can either move upward or downward with some probability. At point $(-1, t_j)$, for example, the variable can attain state $(2, t_{j+1})$ at time t_{j+1} with probability $q^c_{-1}(t_j)$ and state $(-2, t_{j+1})$ at time t_{j+1} with probability $1 - q^c_{-1}(t_j)$.

$$q_i^m(t_j) = \frac{\mu_r(\hat{r}_i(t_j), t_j)\Delta t + \hat{r}_i(t_j) - \hat{r}_{i-J^D(\hat{x}_i(t_j))}(t_{j+1})}{\hat{r}_{i+J^U(\hat{x}_i(t_j))}(t_{j+1}) - \hat{r}_{i-J^D(\hat{x}_i(t_j))}(t_{j+1})} \tag{6.6}$$

If this modified probability still exceeds the desired boundaries because of the maximum and minimum number of jumps possible at each node, simply censor the modified probability to obtain:

$$q_i^c(t_j) = \mathrm{Max}(0, \mathrm{Min}(1, q_i^m(t_j))) \tag{6.7}$$

The binomial tree constructed in the way described above ensures that the discrete-time process of the instantaneous interest rate $\hat{r} = \{\hat{r}(t_i), t_j = j\Delta t, i = 0, \ldots, n\}$ converges in distribution towards the continuous-time process $r=\{r(t), t \quad [0, T]\}$ given by equation (6.1).[4] Given the number of steps n to divide a given time interval, the number of nodes in the tree (which is a useful benchmark or measure to compare with other numerical methods) is simply equal to

$$Q_{NR}(n) = \sum_{i=0}^{i=n}(i+1) = \frac{1}{2}(n+1)(n+2) \tag{6.8}$$

The use of this binomial lattice for the valuation of interest rate derivative securities can be illustrated very clearly by an American call option on a discount bond. Suppose, therefore, that the maturities of the option and the underlying bond are τ_1 and τ_2 respectively. The face value of the bond is 100 and the exercise price of the option is defined as K. The total number of steps to divide the time interval $[0, \tau_2]$ is n_2, whereas the number of steps to divide the subinterval $[0, \tau_1]$ is n_1. The discrete-time value of the bond at each node (i, t_j), with $i=0,\ldots, j$ and $j=0,\ldots, n_2$ is denoted by $\hat{B}_i(t_j)$. The value of the option at each node (i, t_j), with $i=0,\ldots, j$ and $j=0,\ldots, n_1$, finally, is given by $\hat{C}_i(t_j)$.

At maturity of the bond, it is obvious that at each of the corresponding nodes the value of the bond equals its face value, that is:

$$\hat{B}_i(\tau_2) = 100 \tag{6.9}$$

for i=0,…, n_2. The value of the bond at the other nodes can be obtained using the equivalence between the exclusion of arbitrage opportunities and the martingale property of relative values of interest rate dependent securities. The arbitrage-free value of a security is equal to its discounted expectation, where the expectation has to be taken with respect to the unique probability measure under which the relative value of securities are martingales. As the binomial tree is discretely representing the continuous-time process with respect to this unique measure, the following recursive relationship results:

$$\hat{B}_i(t_j) = e^{-\hat{r}_i(t_j)\Delta t}\left(q_i^c(t_j)\hat{B}_{i+1}(t_{j+1}) + (1 - q_i^c(t_j))\hat{B}_{i-1}(t_{j+1})\right) \tag{6.10}$$

with i=0,…, j and i=0,…, n_2−1. These bond values can be used to obtain the value of the call option at maturity of the option,[5] that is:

$$\hat{C}_i(\tau_1) = \mathrm{Max}(\hat{B}_i(\tau_1) - K, 0) \tag{6.11}$$

for i=0,…, n_1. Based on the same recursive relationship,

$$\hat{C}_i(t_j) = e^{-\hat{r}_i(t_j)\Delta t}\left(q_i^c(t_j)\hat{C}_{i+1}(t_{j+1}) + (1 - q_i^c(t_j))\hat{C}_{i-1}(t_{j+1})\right) \tag{6.12}$$

with i=0,…, j and i=0,…, n_1−1. If at some node, premature exercise of the option would be optimal, the value of option determined by discounted expectation at this node can be replaced by its exercise value. In the numerical algorithm this translates into

$$\hat{C}_i(t_j) = \mathrm{Max}(e^{-\hat{r}_i(t_j)\Delta t}(q_i^c(t_j)\hat{C}_{i+1}(t_{j+1}) + (1 - q_i^c(t_j))\hat{C}_{i-1}(t_{j+1})), \hat{B}_i(t_j) - K) \tag{6.13}$$

for i= 0,…, j and i=0,…, n_1−1. The convergence properties of the discrete-time interest rate process translate, in some sense, to the discrete-time option value. If the number of steps n_2, and therefore n_1, increases, the initial value of the option, $\hat{C}_0(0)$, obtained by using the binomial tree, converges to its continuous-time counterpart.[6]

The Method of Tian

The method of Nelson and Ramaswamy (1990) approximates a continuous-time stochastic process by a discrete binomial tree. The binomial probabilities at each node of the tree are chosen such that the instantaneous drift and volatility functions of the stochastic process are matched exactly as the number of steps goes to infinity. At some nodes of the binomial lattice, however, multiple upward and downward jumps may be possible to accomplish this matching and get legitimate probabilities, increasing the numerical complexity of this binomial approximation.

This increased numerical complexity is avoided in the method of Tian (1991) by considering only simple upward and downward jumps and censoring the binomial probability, if necessary. Although this simplification may hurt the numerical accuracy and stability, some of the different steps in this approach are easier to implement and decrease the number of calculations necessary to determine the value of an interest rate derivative security, as will be shown in this section.

The continuous-time process $x=\{x(t), t \in [0, T]\}$ is again approximated by a binomial tree consisting of n steps. Each of these steps represents a subinterval of length $\Delta t = T/n$ and the different values of the approximate process $\hat{x} = \{\hat{x}(t_j), t_j = j\Delta t, j = 0, \ldots, n\}$ at each point of the tree are given by

$$\hat{x}_i(t_j) = (j - 2i)\sqrt{\Delta t} \tag{6.14}$$

with $i=0,\ldots, j$ and $j=0,\ldots, n$. However, the binomial probabilities are now chosen such that the instantaneous drift of the transformed process given by equation (6.2), is matched exactly:[7]

$$q_i(t_j) = \frac{1}{2} + \frac{1}{2}\mu_x(\hat{x}_i(t_j), t_j)\sqrt{\Delta t} \tag{6.15}$$

Similar to the method of Nelson and Ramaswamy, this probability may turn negative or become greater than one because of the particular behavior of the drift and volatility functions. A necessary and sufficient condition to ensure that this probability is legitimate at node (i, t_j) with $i=0,\ldots, j$ and $j=0,\ldots, n-1$ is:[8]

$$|\ \mu_x(\hat{x}_i(t_j), t_j)\ | \leq \frac{1}{\sqrt{\Delta t}} \tag{6.16}$$

The set of nodes for which this condition holds is defined as

$$\mathcal{P} = \left\{(i, t_j), i = 0, \ldots, j; j = 0, \ldots, n-1 \mid \mu_x(\hat{x}_i(t_j), t_j) \,| \leq \frac{1}{\sqrt{\Delta t}}\right\} \tag{6.17}$$

If the initial point of the binomial lattice is an interior point of this sample, the probabilities at the outside regions of the tree are modified to ensure that only elements of $\mathcal{P}$ can be reached. Practically, this means for the lower part of the lattice

$$q_{I^L(t_j)}(t_j) = 1 \quad \text{If } \hat{x}_{I^L(t_j)-1}(t_j) \in \mathcal{P} \text{ and } \hat{x}_{I^L(t_j)}(t_j) \notin \mathcal{P} \tag{6.18}$$

for $j=0,\ldots, n-1$, while for the upper part

$$q_{I^U(t_j)}(t_j) = 0 \quad \text{If } \hat{x}_{I^U(t_j)+1}(t_j) \in \mathcal{P} \text{ and } \hat{x}_{I^U(t_j)}(t_j) \notin \mathcal{P} \tag{6.19}$$

for $j=0,\ldots, n-1$. This inward jumping to modify the branching process is illustrated in Figure 6.2. At each point of the lattice, which is an element of $\mathcal{P}$, the transformed binomial process can either increase or decrease by $\sqrt{\Delta t}$. At point $(-3, t_j)$, which is not an element of $\mathcal{P}$, however, the process can only attain the point $(-2, t_{j+1})$ with probability one. This point $(-2, t_{j+1})$ is again an element of $\mathcal{P}$.

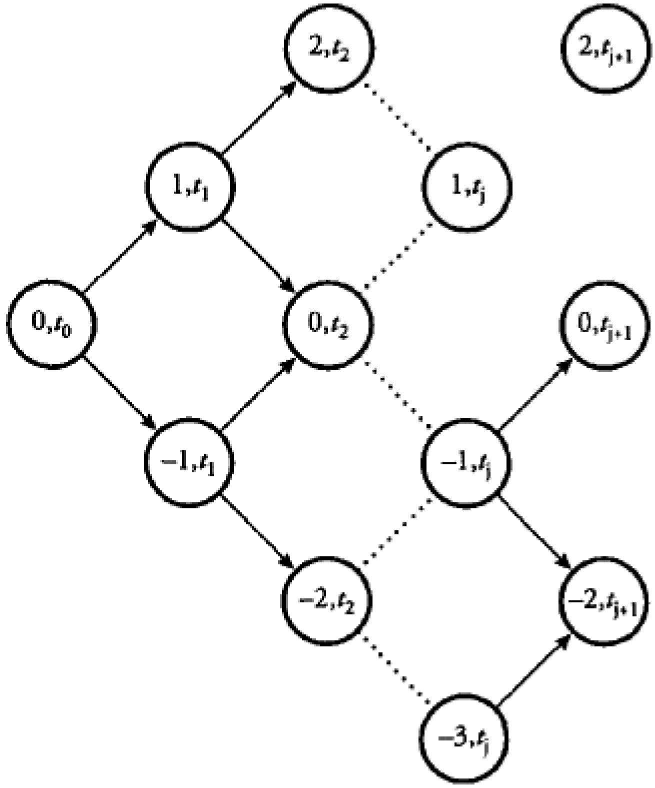

Figure 6.2 The method of Tian

This figure shows the binomial lattice constructed by means of the method of Tian. At any point of the tree, which is an element of $\mathcal{P}$, the variable can either move upward or downward with some probability. At point $(-1, t_1)$, for example, which is an element of $\mathcal{P}$, the variable can attain state $(0, t_2)$ at time t_2 with probability $q_{-1}(t_1)$ or state $(-2, t_2)$ at time t_2 with probability $1-q_{-}$

$_1(t_1)$. At point $(-3, t_j)$, however, which is a point just outside $\mathcal{P}$, the variable can only attain state $(-2, t_{j+1})$ at time t_{j+1} in $\mathcal{P}$, with probability 1.

The binomial tree constructed in this way clearly facilitates a numerical implementation. Multiple upward and downward jumps are avoided and the number of nodes in the tree has decreased. To calculate the actual number of nodes necessary for the numerical illustration of the next section, it will be assumed that $\hat{x}_{I^L(t_j)}(t_j)$ and $\hat{x}_{I^U(t_j)}(t_j)$ are constant for all even and odd j=0,…, n−1. Therefore define

$$K^U = \mathrm{Min}\left(\frac{\hat{x}_{I^U(t_j)}(t_j)}{\Delta t} + 1, n\right) \tag{6.20}$$

$$K^L = \mathrm{Min}\left(\left|\frac{\hat{x}_{I^L(t_j)}(t_j)}{\Delta t}\right| + 1, n\right) \tag{6.21}$$

and

$$K = \mathrm{Min}(K^U, K^L)$$

$$J = \mathrm{Max}(K^U, K^L)$$

The total number of nodes, given the number of steps n, equals[9]

$$Q_{Tian}(n) = \frac{(K+1)(K+2)}{2} + \frac{(J-K)^2}{4} + \tag{6.22}$$

$$(J-K)(K+1) + Ent\left[\frac{(J+K+1)(n-J)}{2}\right]$$

for $J-K$ is even. Otherwise, we have

$$Q_{Tian}(n) = \frac{(K+1)(K+2)}{2} + Ent\left[\frac{(J-K)^2}{4}\right] + \tag{6.23}$$

$$(J-K)(K+1) + \frac{(J+K+1)(n-J)}{2}$$

where Ent[.] denotes the Entier function.

The valuation of interest rate derivative securities by means of this binomial lattice is similar to the approach described in the previous section. The recursive discounted expectation procedure or "backwardation" is even simplified by cutting off the tree, if necessary; thus facilitating numerical implementation.

The Method of Hull and White

In the previous two methods, the continuous-time stochastic process of the instantaneous short-rate is approximated by a discrete binomial process. The approach of Nelson and Ramaswamy (1990) explicitly determines the binomial probabilities at each node of the tree, ensuring convergence in distribution of this discrete process to its continuous-time counterpart. This procedure is modified by Tian (1991) to avoid multiple upward and downward jumps at some nodes and reduce the number of nodes, accordingly.

The method of Hull and White (1990a), however, is originally based on the explicit finite difference method to solve the partial differential equation that every contingent claim obeys. This explicit finite difference method is preferred in comparison to the implicit finite difference method for several reasons. In an extensive numerical comparison of the various difference methods of Geske and Shastri (1985), the explicit finite difference method is concluded to be the most efficient numerical procedure. In addition, this method is computationally much simpler than the implicit method since it does not require the inversion of matrices.

The stochastic process of the transformed underlying state variable $x=\{x(r(t), t), t$ $[0, T]\}$ is approximated by a trinomial tree consisting of n steps. A subinterval of length $\Delta t=T/n$ is spanned by each step of the lattice and the different values of the approximate process $\hat{x} = \{\hat{x}(t_j) t_j = j\Delta t, j = 0, \ldots, n\}$ at each point are given by

$$\hat{x}_i(t_j) = i\Delta x \tag{6.24}$$

for some $\Delta x>0$, $i=-j, -j+1, \ldots, j-1, j$ and $j=0, \ldots, n$. At some point of the lattice (i, t_j), three different points $(i+1, t_{j+1})$, (i, t_{j+1}) and $(i-1, t_{j+1})$ at time t_{j+1} can be reached with probability $q_{i, i+1}(t_j)$, $q_{i,i}(t_{j+1})$ and $q_{i, i-1}(t_j)$, respectively. Figure 6.3 shows the resulting branching process. The trinomial probabilities at each point of the lattice are chosen such that again the instantaneous drift term and volatility functions are matched exactly, yielding[10]

$$q_{i,i+1}(t_j) = \frac{1}{2}\left(\mu_x(\hat{x}, t_j)\frac{\Delta t}{\Delta x} + \frac{\Delta t}{\Delta x^2} + \mu_x(\hat{x}, t_j)^2\left(\frac{\Delta t}{\Delta x}\right)^2\right)$$

$$q_{i,i}(t_j) \quad = 1 - \frac{\Delta t}{\Delta x^2} - \mu_x(\hat{x}, t_j)^2\left(\frac{\Delta t}{\Delta x}\right)^2$$

$$q_{i,i-1}(t_j) = \frac{1}{2}\left(-\mu_x(\hat{x}, t_j)\frac{\Delta t}{\Delta x} + \frac{\Delta t}{\Delta x^2} + \mu_x(\hat{x}, t_j)^2\left(\frac{\Delta t}{\Delta x}\right)^2\right)$$

These probabilities are positive if both

$$\frac{\Delta t}{\Delta x^2} + \mu_x(\hat{x}, t_j)^2\left(\frac{\Delta t}{\Delta x}\right)^2 < 1 \tag{6.25}$$

and

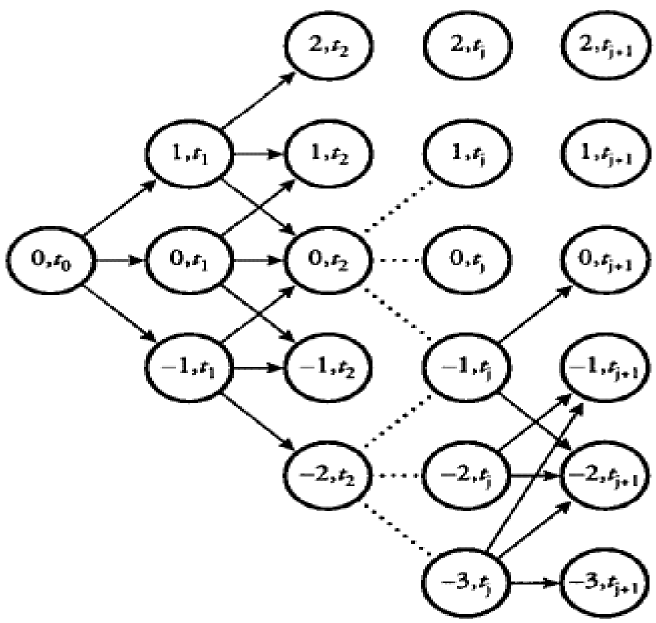

Figure 6.3 The method of Hull and White

This figure shows the trinomial lattice constructed by means of the method of Hull and White. At any point of the tree, which is an element of $\mathcal{P}$, the variable can move upward or downward or stay the same with some probability. At point $(-1, t_1)$, for example, which is an element of the variable can attain state $(0, t_2)$ $(-1, t_2)$ or $(-2, t_2)$ at time t_2 with probabilities $q_{-1,0}(t_1)$, $q_{-1,-1}(t_1)$ and $q_{-1,-2}(t_1)$, respectively. At point $(-3, t_j)$, however, which is a point just outside $\mathcal{P}$, the variable can only attain states $(-1, t_{j+1})$, $(-2, t_{j+1})$ or $(-3, t_{j+1})$ at time t_{j+1} in $\mathcal{P}$ with probabilities $q_{-3,-1}(t_j)$, $q_{-3,-2}(t_j)$ amd $q_{-3,-3}(t_j)$, respectively.

$$\mu_x(\hat{x}, t_j)\frac{\Delta t}{\Delta x} < \frac{\Delta t}{\Delta x^2} + \mu_x(\hat{x}, t_j)^2\left(\frac{\Delta t}{\Delta x}\right)^2 \qquad (6.26)$$

are satisfied. Again, the sample of nodes fullfilling these conditions is defined as

$$\mathcal{P} = \{(i, t_j), i = -j, \ldots, j; j = 0, \ldots, n \mid (6.25) \text{ and } (6.26)\} \qquad (6.27)$$

It is clearly possible to let $\mathcal{P}$ contain all nodes if the instantaneous drift term $\mu_x(x, t)$ is bounded.[11]

However, due to a possible singularity of the volatility function of the stochastic process of the short rate, just as with the one discussed in the previous two sections, some points of the original tree may not be elements of $\mathcal{P}$. Modification of the jump process in these particular nodes, similar to the binomial lattice approaches, enables retention of positive probabilities and ensures convergence. Suppose therefore, for example, that at some point (i, t_j) of the lattice, one wants to reach the following three different points $(i+k+1, t_{j+1})$, $(i+k, t_{j+1})$ and $(i+k-1, t_{j+1})$ at time t_{j+1} with probability $q_{i,i+k+1}(t_j)$, $q_{i,i+k}(t_{j+1})$ and $q_{i,i+k-1}(t_j)$, respectively. Matching of the first and second moments now yields:

$$q_{i,i+k+1}(t_j) = \frac{1}{2}\left((k^2-k) + (1-2k)\mu_x(\hat{x}, t_j)\frac{\Delta t}{\Delta x} + \frac{\Delta t}{\Delta x^2} + \mu_x(\hat{x}, t_j)^2\left(\frac{\Delta t}{\Delta x}\right)^2\right)$$

$$q_{i,i+k}(t_j) \quad = 1 - k^2 + 2k\mu_x(\hat{x}, t_j)\frac{\Delta t}{\Delta x} - \frac{\Delta t}{\Delta x^2} - \mu_x(\hat{x}, t_j)^2\left(\frac{\Delta t}{\Delta x}\right)^2$$

$$q_{i,i+k-1}(t_j) = \frac{1}{2}\left((k^2+k) - (1+2k)\mu_x(\hat{x}, t_j)\frac{\Delta t}{\Delta x} + \frac{\Delta t}{\Delta x^2} + \mu_x(\hat{x}, t_j)^2\left(\frac{\Delta t}{\Delta x}\right)^2\right)$$

If $\hat{x}_{I^L}(t_j)$ is the smallest negative value of the approximate discrete process at time t_j, j=0,…, n, which is not an element of $\mathcal{P}$, formally identified as

$$\hat{x}_{I^L+1}(t_j) \in \mathcal{P} \quad \text{And} \quad \hat{x}_{I^L}(t_j) \notin \mathcal{P} \tag{6.28}$$

the number of extra upward jumps $K(I^L)$ at this point necessary to return to the trinomial tree is:

$$K(I^L) = \begin{cases} \text{The smallest, positive integer } k \text{ such that} \\ \begin{pmatrix} q_{I^L, I^L+K(I^L)+1}(t_j) \\ q_{I^L, I^L+K(I^L)}(t_j) \\ q_{I^L, I^L+K(I^L)-1}(t_j) \end{pmatrix} \geq 0 \quad \text{If } \hat{x}_{I^L+K(I^L)+1}(t_{j+1}) \in \mathcal{P} \\ 1 \qquad\qquad\qquad\qquad\qquad\qquad \text{Otherwise} \end{cases}$$

This modified upward jumping is illustrated in Figure 6.3. At point (−3, t_j), which is just outside $\mathcal{P}$, the process can attain the nodes (−l, t_{j+1}), (−2, t_{j+1}) and (−3, t_{j+1}) with probabilities $q_{-3,-1}(t_j)$, $q_{-3,-2}(t_j)$ and $q_{-3,-3}(t_j)$, respectively. The value of $K(I^L)$ in this case is equal to 1.

Similarly, if $\hat{x}_{I^U}(t_j)$ is the smallest positive value of the approximate discrete process at time t_j, j=0,…, n, which is not an element of $\mathcal{P}$, formally identified as

$$\hat{x}_{I^U-1}(t_j) \in \mathcal{P} \quad \text{and} \quad \hat{x}_{I^U}(t_j) \notin \mathcal{P} \tag{6.29}$$

the number of extra downward jumps $K(I^U)$, at this point, necessary to return to the trinomial tree is:

$$K(I^U) = \begin{cases} \text{The smallest, negative integer } k \text{ such that} \\ \begin{pmatrix} q_{I^U, I^U+K(I^U)+1}(t_j) \\ q_{I^U, I^U+K(I^U)}(t_j) \\ q_{I^U, I^U+K(I^U)-1}(t_j) \end{pmatrix} \geq 0 \quad \text{If } \hat{x}_{I^U+K(I^U)-1}(t_{j+1}) \in \mathcal{P} \\ -1 \qquad\qquad\qquad\qquad\qquad \text{Otherwise} \end{cases}$$

If at these points one of the trinomial probabilities is still negative, the three probabilities are censored by equating them to 1/3

To calculate the number of nodes in the trinomial tree, again the assumption is made that $\hat{x}_{I^U}(t_j)$ and $\hat{x}_{I^L}(t_j)$ are constant for all j=0, …, n. If the number of steps in the lattice is n, the number of nodes Q is[12]

$$Q_{HW}(n) = (K+1)^2 + \frac{(J-K)(J+3K+3)}{2} + (J+K+1)(n-J) \qquad (6.30)$$

with

$$K^U = \mathrm{Min}\left(\frac{\hat{x}_{I^U(t_j)}(t_j)}{\Delta t} + 1, n\right) \qquad (6.31)$$

$$K^L = \mathrm{Min}\left(\left|\frac{\hat{x}_{I^L(t_j)}(t_j)}{\Delta t}\right| + 1, n\right) \qquad (6.32)$$

and

$$K = \mathrm{Min}(K^U, K^L)$$

$$J = \mathrm{Max}(K^U, K^L)$$

As explained and illustrated in the section on the method of Nelson and Ramaswamy (pp. 112–5 above), the value of an interest rate derivative security is determined by calculating its discounted expectation. Starting at maturity of the claim, the value of the short-term rate of interest at each node, together with the corresponding trinomial probabilities, are used recursively to obtain this expectation. In the case of an American option, the early-exercise feature can simply be incorporated by replacing the option's value, determined recursively, by the early-exercise value. Because the discrete-time process of the interest rate converges in distribution to its continuous-time counterparts as the number of steps increases, convergence of option values, obtained numerically, is easily shown.[13]

COMPARING PATH-INDEPENDENT NUMERICAL METHODS: THE MODEL OF COX, INGERSOLL AND ROSS

This section illustrates and compares the three numerical methods by valuing options on discount bonds using the model of Cox, Ingersoll and Ross (1985) (CIR), discussed in Chapter 5, pp. 76–83. Because this model incorporates a lot of empirical characteristics of interest rates and, as such, is the subject of extensive empirical research, it is interesting to show an actual implementation of the numerical methods using this particular model. In addition, an attempt will be made to develop some decision rules to compare different numerical methods, and this model serves as a nice benchmark to illustrate this analysis.

This section starts with a derivation of the transformed process using the CIR model, necessary for the application of the three different numerical methods. The instantaneous short-term rate of interest in this model obeys the following stochastic differential equation in the risk-neutral economy:

$$dr(t) = \kappa(\theta - r(t))dt + \sigma\sqrt{r(t)}d\tilde{W}(t) \tag{6.33}$$

The parameters κ, θ and σ are positive, real-valued constants, while the stochastic process $\tilde{W} = \{\tilde{W}(t), t \in [0, T]\}$ is a Standard Brownian Motion initialized at zero. The initial value of the interest rate, $r(0)$, is a known, real-valued constant.

The transformation used to obtain a stochastic process with a unit-instantaneous volatility function is, according to equation (6.3),

$$x(r, t) = \int_{r(0)}^{r(t)} \frac{dz}{\sigma_r(z, t)} \tag{6.34}$$

yielding

$$x(r, t) = \int_{r(0)}^{r(t)} \frac{dz}{\sigma\sqrt{z}} = \frac{2\sqrt{r(t)}}{\sigma} - \frac{2\sqrt{r(0)}}{\sigma} \qquad 6.35$$

The resulting stochastic differential equation of the transformed variable equals

$$dx(t) = \left(\frac{\alpha_1}{\alpha_2 + \alpha_3 x(t)} + \alpha_4 x(t) + \alpha_5\right)dt + d\tilde{W}(t) \tag{6.36}$$

with

$$\alpha_1 = 4\kappa\theta - \sigma^2 \qquad \alpha_4 = -\frac{\kappa}{2}$$

$$\alpha_2 = 4\sigma\sqrt{r(0)} \qquad \alpha_5 = -\frac{\kappa}{\sigma}\sqrt{r(0)}$$

$$\alpha_3 = 2\sigma^2$$

Given the characteristics of the contingent claim and the particular values of the parameters of the stochastic process, interest rate contingent claims can be valued according to the different numerical procedures described and discussed in the previous section. In the case of a call option on a bond, the lattice has to span the maturity of the option only and not the maturity of the underlying security as well. At each of the final nodes, the corresponding instantaneous interest rate in that state is known and because the value of the underlying bond is an explicit analytic function of this rate, the exercise value of the option is readily obtained.[14]

An interesting problem, now, is the way in which these procedures have to be compared with respect to convergence and stability. In a number of papers, in which actual comparisons between some numerical procedures are performed,[15] a contingent claim, for which a closed-form solution exists, serves as a benchmark. This claim is then evaluated for different claim characteristics, parameter values and an increasing number of steps. Based on this increasing number of steps, a judgement can be made about the convergence and stability of the various numerical methods.

In the event that numerical methods have to be used because of the complex characteristics of the derivative security in question, it seems reasonable to use a claim as benchmark for which a closed-form solution does exist. For different characteristics of this benchmark, a sufficient number of steps can be determined to guarantee the desired numerical properties such as stability and convergence. However, as will be clear from the discussion of the term structure models so far, this closed-form contingent claim does not always exist and a comparison of the different methods on the basis of an increasing number of steps is difficult.

Even if closed-form solutions for some claims do exist, the different claim values, determined numerically, may exhibit very typical oscillatory patterns. Showing the claim's values, therefore, for some specific number of steps is definitely not sufficient to examine the numerical properties. To illustrate this statement, Figure 6.4 shows the numerical value of a 10 per cent in-the-money European call option on a discount bond for different numbers of steps. The maturity of the option and the underlying bond are five and ten years, respectively, whereas the face value of the bond is 100. The option values are obtained by using the method of Hull and White (1990a).

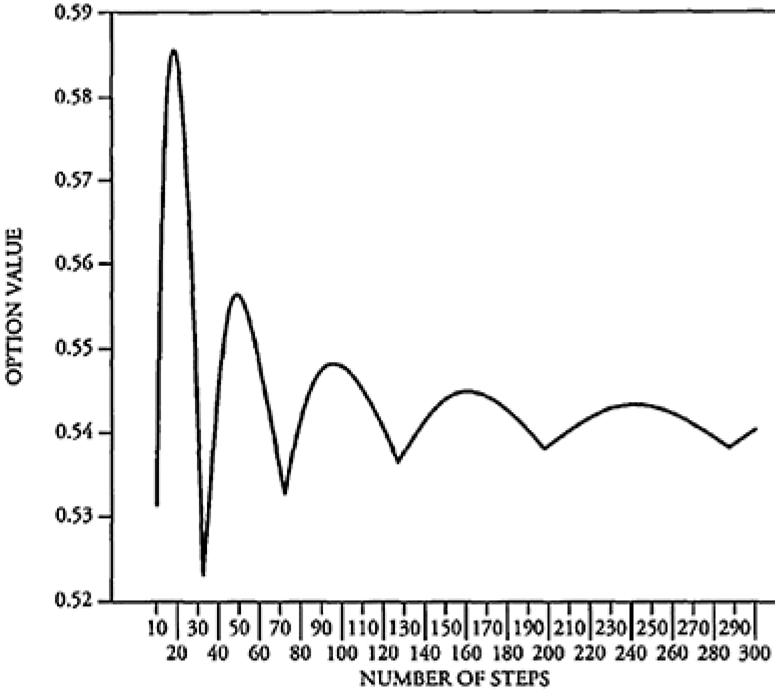

Figure 6.4 Numerically obtained option values

This figure shows numerically obtained European call option values as a function of the number of steps. The numerical approach used is the method of Hull and White (1990a). The maturity of the option is five years and the exercise price of the option is 1.1 times the forward price of the underlying bond. The maturity of the underlying bond is ten years, while the face value is 100. The values of the parameters $r(0)$, κ, θ and σ are 0.10, 0.2, 0.10 and 0.10, respectively.

It is obvious from this specific oscillatory pattern of the option values that a simple examination of convergence is not possible. In observing a particular array of increasing number of steps, overestimating or under-estimating the numerical accuracy, and thereby the rate of convergence, is almost unavoidable.

To take these typical patterns of option values into consideration and to be able to compare the numerical properties of interest rate processes for which closed-form option values do not exist, this section proposes the following procedure, using the CIR interest rate model and the numerical procedures, mentioned above, to illustrate it.

Convergence properties of the different approaches can be compared by valuing European call and put options on discount bonds. Although the values of these contingent claims allow for closed-form solutions, we do not use this property, which will keep the presentation as general as

possible. For a specific scenario of parameter values and option characteristics, the benchmark value is determined as follows:

$$\hat{C}_{Bench}(n_{Bench}) = \frac{\hat{C}_{NR}(n_{Bench}) + \hat{C}_{Tian}(n_{Bench}) + \hat{C}_{HW}(n_{Bench})}{3} \quad (6.37)$$

where $\hat{C}_{NR}(n)$, $\hat{C}_{Tian}(n)$ and $\hat{C}_{HW}(n)$ denote the numerically obtained option values based on n steps using the three methods, respectively. The number of steps, n_{Bench}, to obtain the benchmark value is, for each of the scenarios, equal to 300, although even this number, of course, may introduce some error due to the above-mentioned cyclical behavior of option values.[16]

Given a particular scenario and the corresponding benchmark value, the maximum number of steps n_{NR}, n_{Tian} and n_{HW} for which the difference between the numerical and benchmark option value is greater than one cent can be recursively determined. More formally, this means

$$| \hat{C}_{Bench}(n_{Bench}) - \hat{C}_{NR}(n_{NR}) | \geq 0.01$$

$$| \hat{C}_{Bench}(n_{Bench}) - \hat{C}_{Tian}(n_{Tian}) | \geq 0.01$$

$$| \hat{C}_{Bench}(n_{Bench}) - \hat{C}_{HW}(n_{HW}) | \geq 0.01$$

and

$$| \hat{C}_{Bench}(n_{Bench}) - \hat{C}_{NR}(n) | \leq 0.01 \forall n > n_{NR}$$

$$| \hat{C}_{Bench}(n_{Bench}) - \hat{C}_{Tian}(n) | \leq 0.01 \forall n > n_{Tian}$$

$$| \hat{C}_{Bench}(n_{Bench}) - \hat{C}_{HW}(n) | \leq 0.01 \forall n > n_{HW}$$

Table 6.1 Scenarios used for numerical comparison

	Low case	*Base case*	*High case*
$r(0)$	0.06	0.10	0.14
κ	0.2	0.2	0.2
θ	0.1	0.1	0.1
σ	0.1	0.1	0.1
α	−0.1	0	0.1
τ_0	2.5	5	7.5
τ_B	10	10	10

This table contains the different scenarios used to compare the numerical methods using the CIR model. The parameters κ, θ, σ and the initial value of the interest rate $r(0)$ describe the stochastic dynamics of the interest rate, as is discussed in Section 5.1. The symbols τ_0 and τ_B denote the maturity of the option and the underlying bond in years, respectively. Finally, the parameter α represents the degree by which the option is in-the-money or out-of-the-money. In the case that this parameter is zero, the exercise price of the option is equal to the forward price of the underlying bond.

Based on these numbers n_{NR}, n_{Tian} and n_{HW}, then, the computational efficiency to achieve this rate of convergence is measured by the number of nodes of the lattice.

The various scenarios used to perform this numerical comparison can be classified according to the particular values of the stochastic interest rate process and the dififerent option characteristics, such as time-to-maturity, put or call and the degree to which the option is in-the-money or outof-the-money. Preferable is a low case, base case and high case, for each of these values and characteristics, resulting in $3^7 \times 2=4374$ different scenarios. This thesis, however, will only compare the numerical approaches by means of the following cases shown in Table 6.1, resulting in 54 scenarios. These scenarios cover a range of reasonable parameter values for which it is interesting to examine the convergence of the different approaches.

The results shown in Table 6.2 indicate that the binomial lattice approach of Tian is generally the most efficient numerical procedure to value options on bonds using the CIR model. The average number of nodes $Q_{Tian}(n_{Tian})$ for which the corresponding option value is at least within one cent of the benchmark value, is 3140, compared to 3743 and 4298 for the method of Hull and White and Nelson and Ramaswamy. In addition, the corresponding standard deviation of this number is also significantly lower. It is further interesting to note that in a comparison of the methods with respect to the mean and standard deviation of the number of steps, the trinomial approach would be preferred, while the two binomial methods would be almost indistinguishable. The decrease in the number of nodes in the method of Tian because of the mean reversion and the resulting inward jumps, significantly increases the computational efficiency.

Table 6.2 Statistics of numerical comparison

	Nelson and Ramaswamy		*Tian*		*Hull and White*	
	n_{NR}	$Q_{NR}(n_{NR})$	n_{Tian}	$Q_{Tian}(n_{Tain})$	n_{Hw}	$Q_{HW}(n_{HW})$
Mean	87	4,298	88	3,140	69	3,743
St. Dev.	3.85	370	3.96	241	4.06	375
Maximum	156	12,403	148	7,269	149	13,320
Minimum	21	253	21	208	12	133

This table contains the summary statistics of the comparison of the three numerical procedures. The different scenarios are described in Table 6.1, and the total number is 54. The numbers n_{NR}, n_{Tian} and n_{Hw} represent the number of steps for which convergence within one cent of the benchmark value has been obtained. The functions $Q_{NR}(n_{NR})$ $Q_{Tian}(n_{Tian})$ and $Q_{HW}(n_{HW})$ give the corresponding number of nodes, useful as a measure of computational efficiency.

THE PATH-INDEPENDENT INTEREST RATE MODEL OF HEATH, JARROW AND MORTON I

This section illustrates a two-factor model of Heath, Jarrow and Morton (1990a, 1992) I, discussed in Section 5.2.[17] Although their original discrete approximation of the continuous-time interest rate model is path-dependent, this section will derive a numerical procedure that is pathindependent.[18] In addition, an extension of the interest rate tree will be presented to determine any kind of desired hedge ratios, which are useful for the management of interest rate risk. At the end of this section, various features and numerical efficiency of this method will be illustrated.

The instantaneous short-term rate of interest is a function of two independent factors and a term, ensuring that a given term structure of interest rates at the valuation date is obtained in the risk-neutral economy,[19] that is,

$$r(t) = f(0, t) + h_1(t) + h_2(t) + \sigma_1 x_1(t) + \sigma_2 x_2(t) \tag{6.38}$$

with

$$h_1(t) = \frac{1}{2}\sigma_1^2 t^2 \tag{6.39}$$

$$h_2(t) = \sigma_2^2 \frac{2}{\kappa^2}(1 - e^{-\frac{\kappa}{2}t})^2 \tag{6.40}$$

and

$$x_1(t) = \tilde{W}_1(t) \tag{6.41}$$

$$x_2(t) = \int_0^t e^{-\frac{\kappa}{2}(t-s)} d\tilde{W}_2(t) \tag{6.42}$$

In these expressions, $f(0, t)$ is the exogenous term structure of forward rates at the initial trade date 0. The parameters σ_1, κ and σ_2 are real-valued constants and the processes $\tilde{W}_1 = \{\tilde{W}_1(t), t \in [0, T]\}$ and $\tilde{W}_2 = \{\tilde{W}_2(t), t \in [0, T]\}$ are independent Standard Brownian Motions initialized at zero.

According to Arnold (1974, p. 134), the factor $x_2(t)$ is a unique solution to an Ornstein—Uhlenbeck process and the stochastic process of the two factors are therefore:

$$dx_1(t) = d\tilde{W}_1(t) \tag{6.43}$$

$$dx_2(t) = -\frac{\kappa}{2} x_2(t)dt + d\tilde{W}_2(t) \tag{6.44}$$

with $x_1(0)=0$ and $x_2(0)=0$. The method of Tian (1991) can be employed to construct a tree with maturity T and the number of steps n. The two continuous-time stochastic factors $x_1(t)$ and $x_2(t)$ are approximated in this method by two binomial factors $\hat{x}_1(t)$ and $\hat{x}_2(t)$. Starting with the initial values $\hat{x}_1(0) = 0$ and $\hat{x}_2(0) = 0$, at each point of the lattice each binomial factor $\hat{x}_1(t_i)$ and $\hat{x}_2(t_i)$ can either increase or decrease by $\sqrt{\Delta t}$ at time t_i, where $\Delta t=T/n$ and $t_i= i\Delta t$, $i=0,\ldots, n-1$, resulting in four possible states. In case of the second factor, given its value $\hat{x}_2(t)$ at time t, the values $\hat{x}_2(t) + \sqrt{\Delta t}$ and $\hat{x}_2(t) - \sqrt{\Delta t}$ at time $t+\Delta t$ can be reached with probability $\frac{1}{2} - \frac{\kappa}{4}\hat{x}\ (t)\sqrt{\Delta t}$ and $\frac{1}{2} + \frac{\kappa}{4}\hat{x}_2(t)\sqrt{\Delta t}$, respectively. If one of these probabilities turns out to be negative at some point of the lattice, a simple inward jump with probability one is modelled. It is obvious that in case of the first factor, the corresponding values can be reached with equal probability. At each point of this two-dimensional tree, finally, the value of the instantaneous interest rate $r(t)$ is then obtained from equation (6.38) and interest rate contingent claims are valued by means of the familiar recursive computation, starting at maturity of this claim.

Similar to the CIR model (1985) in the previous section, the value of a discount bond, on which the claim may be dependent at maturity, is an explicit analytic function of the short-rate at this time. This known formula considerably enhances the computational speed of the algorithm, because the interest rate tree has only to span the maturity of the claim instead of the maturity of the underlying value.[20] At time t, therefore, the value of a zero-coupon $P(x_1(t), x_2(t), t, \tau)$ bond with remaining time-to-maturity τ, is given the values of the two factors $x_1(t)$ and $x_2(t)$ at this time,[21]

$$P(x_1(t), x_2(t), t, \tau) = e^{-\int_t^{t+\tau} f(0,s)ds} P_1(x_1(t), t, \tau) P_2(x_2(t), t, \tau) \tag{6.45}$$

where

$$P_1(x_1(t), t, \tau) = exp(-\tau\sigma_1 x_1(t) + V_1(t, \tau)) \tag{6.46}$$

with

$$V_1(t, \tau) = -\frac{1}{2}\sigma_1^2 t\tau(t + \tau) \tag{6.47}$$

and

$$P_2(x_2(t), t, \tau) = exp\left(-\frac{2}{\kappa}(1 - e^{\frac{\kappa}{2}\tau})\sigma_2 x_2(t) + V_2(t, \tau)\right) \tag{6.48}$$

with

$$V_2(t, \tau) = 2\sigma_2^2 \frac{1 + (2 - e^{-\frac{\kappa}{2}(t+\tau)})^2 - (2 - e^{-\frac{\kappa}{2}\tau})^2 - (2 - e^{-\frac{\kappa}{2}t})^2}{\kappa^3} \tag{6.49}$$

Because the first factor in this model is similar to the model of Ho and Lee (1986) and changes in this factor result in parallel shifts in the yield curve, the derivative of the bond $P(x_1(t), x_2(t), t, \tau)$ with respect to this factor equals the Redington (1952) duration[22] D_P,

$$D_P = -\frac{(dP(x_1(t), x_2(t), t, \tau)/P(x_1(t), x_2(t), t, \tau)}{dx_1} = \tau \tag{6.50}$$

Changes in the value of the second factor can now be interpreted as causing a rotation R_P of the term structure of interest rates,

$$R_P = -\frac{(\partial P(x_1(t), x_2(t), t, \tau)/P(x_1(t), x_2(t), t, \tau)}{\partial x_2} = \frac{1 - e^{-\frac{\kappa}{2}\tau}}{\frac{\kappa}{2}} \tag{6.51}$$

Although the sensitivities of discount bonds with respect to changes in the underlying two factors can be derived analytically, it would be interesting to be able to obtain these sensitivities for any interest rate contingent claim. To obtain these values numerically, the interest rate tree described above can be expanded in the following way.[23] The interest rate tree will be increased by two time steps. The new valuation date starts with an exogenous term structure $R^*(0, \tau)$ such that after two steps the term structure $R^*(0, 0, 2\Delta t, \tau)$ resulting from an up and down move of each of the factors is equal to the exogenous term structure $R(0, \tau)$ at the original initial valuation date 0, that is,

$$R^*(0, \tau + 2\Delta t) = \frac{\tau R(0, \tau) + 2\Delta t R(0, 2\Delta t) + V_1(2\Delta t, \tau + 2\Delta t) + V_2(2\Delta t, \tau + 2\Delta t)}{\tau + 2\Delta t} \tag{6.52}$$

The values of the contingent claim at time $t=2\Delta t$ at two up and down moves can then be used to determine the sensitivity of this claim with respect to the underlying factors. In the case of the discount bond, the numerical values of duration and rotation are

$$\hat{D}_P = -\frac{(P(2\sqrt{\Delta t}, 0, 2\Delta t, \tau + 2\Delta t) - P(-2\sqrt{\Delta t}, 0, 2\Delta t, \tau + 2\Delta t))}{P(0, 0, 2\Delta t, \tau + 2\Delta t)} / 4\sqrt{\Delta t} \tag{6.53}$$

and

$$\hat{R}_P = -\frac{(P(0, 2\sqrt{\Delta t}, 2\Delta t, \tau + 2\Delta t) - P(0, -2\sqrt{\Delta t}, 2\Delta t, \tau + 2\Delta t))}{P(0, 0, 2\Delta t, \tau + 2\Delta t)} / 4\sqrt{\Delta t} \tag{6.54}$$

To illustrate the proposed numerical method, start by calculating the initial value, duration and rotation of a zero-coupon bond with a maturity of ten years and face value of 100. The exogenous term structure at the initial trade date is assumed to be flat at 10 per cent, while the parameters σ_1, κ and σ_2 are 0.02, 0.4 and 0.02, respectively. The numerically obtained results are shown in Table 6.3.

Although this example is only expository, it can be seen that convergence within one cent of the exact value is obtained after more than thirty steps. The rate of convergence of duration and rotation, however, seems to be even higher.

To extend the example, the value of a European call option will also be computed. The value of this claim $C(0, \tau_1, \tau_2, K)$ with maturity τ_1, exercise price K and written on a discount bond with maturity τ_2 at the initial valuation date is:[24]

$$C(0, \tau_1, \tau_2, K) = P(0, 0, 0, \tau_2)N(d_1) - KP(0, 0, 0, \tau_1)N(d_2) \tag{6.55}$$

with

$$d_1 = \ln\left(\frac{P(0, 0, 0, \tau_2)}{P(0, 0, 0, \tau_1)}K\right) \Big/ v + \frac{1}{2}v$$

$$d_2 = d_1 - v$$

$$v^2 = \sigma_1(\tau_2 - \tau_1)^2\tau_1 +$$

$$\frac{4\sigma_2^2}{\kappa^3}\left(\left(1 - e^{-\frac{\kappa}{2}(\tau_2 - \tau_1)}\right)^2 - \left(e^{-\frac{\kappa}{2}\tau_2} - e^{-\frac{\kappa}{2}\tau_1}\right)^2\right)$$

Table 6.3 Numerical bond values

Number of steps	*Value*	D_p	R_p
10	36.7870	10.0534	4.3334
20	36.7862	10.0267	4.3283
30	36.7865	10.0178	4.3267
40	36.7868	10.0133	4.3258
50	36.7870	10.0107	4.3253
60	36.7871	10.0089	4.3250
Exact value	36.7879	10.0000	4.3233

This table contains the numerical values of a discount bond with a maturity of ten years, its duration and rotation for a different number of steps using the model of Heath, Jarrow and Morton I. The initial term structure is flat at 10 per cent, while the values of the parameters σ_1, κ and σ_2 are 0.02, 0.4 and 0.02, respectively.

Table 6.4 Numerical option values

Number of steps	*Value*	$\frac{\partial C}{\partial P(\tau_2)}$	$\frac{\partial C}{\partial P(\tau_1)}$
10	7.6301	0.8183	−0.2711
20	7.6202	0.8220	−0.2743
30	7.6161	0.8234	−0.2755
40	7.6201	0.8238	−0.2760
50	7.6152	0.8246	−0.2766
60	7.6186	0.8246	−0.2767
Exact value	7.6156	0.8262	−0.2782

This table contains the numerical values of a European call option with a maturity of two years written on a discount bond with a maturity of ten years for a different number of steps using the model of Heath, Jarrow and Morton I. The exercise price of the option is equal to the current value of the underlying bond. The initial term structure is flat at 10 per cent, while the values of the parameters σ_1, κ and σ_2 are 0.02, 0.4 and 0.02, respectively.

The extension of the two-dimensional interest rate tree can also be applied to obtain the hedge ratios of this option with respect to the underlying bond and the discount bond with a maturity equal to the maturity of the option. For this purpose, the following relationship between hedge ratios and sensitivities with respect to the underlying factors will be used:

$$\begin{pmatrix} \dfrac{\partial C}{\partial P(\tau_1)} \\ \dfrac{\partial C}{\partial P(\tau_2)} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial P(\tau_1)}{\partial x_1} & \dfrac{\partial P(\tau_2)}{\partial x_1} \\ \dfrac{\partial P(\tau_1)}{\partial x_2} & \dfrac{\partial P(\tau_2)}{\partial x_2} \end{pmatrix}^{-1} \begin{pmatrix} \dfrac{\partial C}{\partial x_1} \\ \dfrac{\partial C}{\partial x_2} \end{pmatrix} \tag{6.56}$$

In Table 6.4, the numerical results are shown for the same parameter values and different number of steps. The exercise price of the option is equal to the current value of the underlying bond, while the maturity of the option and the underlying bond are two and ten years, respectively.

Convergence within one cent of the exact value has now been achieved after forty steps. Moreover, it can be seen from this table that extending the interest rate tree to obtain the hedge ratios also yields accurate results.

THE PATH-DEPENDENT INTEREST RATE MODEL OF HEATH, JARROW AND MORTON II

Path-dependent interest rate trees are characterized by the simple property that each of the nodes in the tree can be reached by one particular node with some positive probability. The previous section defined some constraints on the stochastic interest rate process such that it is possible to construct lattices in which the different paths recombine, making the tree path-independent. As was noted in the introduction to this chapter, it is tempting to consider a path-independent interest rate tree numerically to be more attractive than a path-dependent tree. The number of nodes in the tree as a function of the number of steps increases less than exponentially when the paths are recombining, resulting in the possibility of evaluating more steps in practical applications.

Recombining interest rate trees, however, imposes some constraints on the values of the interest rate at each of the nodes and the corresponding probabilities, which may decrease the rate of convergence. In addition, the class of interest rate processes for which it is possible to construct these trees is not general enough to be able to deal with all the interest rate models discussed in Chapter 5. As a result, it is necessary and interesting to pay some attention to those models for which only path-dependent interest trees can be constructed to value contingent claims.

This section will discuss the construction of a path-dependent interest rate tree using the model of Heath, Jarrow and Morton (1990b) II for two reasons. First, this model serves as a nice example of an interest rate process for which only a path-dependent tree can be constructed. Second, the multi-factor extension of this model is used as an example in Chapter 9 to illustrate some serious limitations of a commonly used estimation technique when valuing interest rate derivative securities.

In Chapter 5 we derived the continuous-time stochastic process of the forward rate curve in the risk-neutral economy:

$$f(t,s)=f(0,s)+\int_0^t \sigma(v,s,f(v,s))\int_v^s \sigma(v,w,f(v,w))dwdv+ \tag{6.57}$$

$$\int_0^t \sigma(v,s,f(v,s))d\tilde{W}(v)$$

In this expression, $f(0, t)$, t $[0, T]$ is the familiar exogenous forward rate curve at the initial valuation date, whereas the process $\tilde{W}=\{\tilde{W}(t)_t$ $[0, T]\}$ represents a Standard Brownian Motion initialized at zero. The general volatility function $\sigma(t, s, f(t, s))$, t, s $[0, T]$, $s > t$, is dependent on the calendar time, the maturity of the forward rate and the value of the forward rate itself. Although the discussion of this model in Chapter 5 is mainly concerned with this general specification, the particular volatility function used in the empirical analysis of Heath *et al.* (1990b), and used in Chapter 9 to illustrate some of the estimation problems with this model, is equal to

$$\sigma(t,s,f(t,s))=\sigma Min(f_{Max},f(t,s)) \tag{6.58}$$

where σ and f_{Max} are positive, real-valued constants.

The continuous-time stochastic process $\tilde{W}=\{\tilde{W}(t), t\in[0,T]\}$ is approximated by a path-dependent tree consisting of n steps. A subinterval of length $\Delta t = T/n$ is spanned by each step of the of the tree the approximate discrete process $\hat{W}=\{\hat{W}(t_j), t_j=j\Delta t$, $j=0,\dots, n\}$ can either increase or decrease by $\sqrt{\Delta t}$.

The continuous-time stochastic process of the forward rate curve can now be approximated by the following discrete-time process $\hat{f}=\{\hat{f}(t_j,s), t_j=j\Delta t, j=0,\dots,n, s\in[t_{j+1},T]\}$. At some time t_j, the forward rate can attain two values at time t_{j+1} with equal probability due to the binomial approximation of the Standard Brownian Motion, that is:

$$\hat{f}(t_{j+1},s)=\begin{cases}\hat{f}(t_j,s)+\sigma(t_j,s)\sqrt{\Delta t}+\alpha(t_j,s)\Delta t\\ \hat{f}(t_j,s)-\sigma(t_j,s)\sqrt{\Delta t}+\alpha(t_j,s)\Delta t\end{cases}$$

with $j=0,\dots, n-1$, s $[t_{j+1}, T]$. For ease of exposition, the dependence of the different functions on the current value of the forward rate has been suppressed. The resulting interest rate tree is path-dependent and at time t_j, $j = 0,\dots, n$, the number of different states is 2^j.

The analytic expression of the discrete function $\alpha(t_j, s)$ can readily be obtained from the continuous-time stochastic process given by equation (6.57). As the number of steps n increases, the resulting discrete process converges to its continuous-time counterpart. However, the path-dependent tree limits the number of steps to be evaluated because the number of nodes grows exponentially. To increase convergence, therefore, the function

$\alpha(t_j, s)$ is chosen such that the exogenous term structure of forward rates at the initial valuation date is always obtained, whatever the number of steps *n*, that is:

$$P(t_j, s - t_j) = P(t_j, \Delta t) E_Q(P(t_{j+1}, s - t_{j+1}) \mid \mathcal{F}_{t_j}) \tag{6.59}$$

for $j = 0, \ldots, n$, $s \quad [t_j, T]$. This expression can be rewritten as[25]

$$exp\left(-\int_{t_{j+1}}^{s} f(t_j, u)du\right) = \frac{1}{2} exp\left(-\int_{t_{j+1}}^{s} [f(t_j, u) + \alpha(t_j, u)\Delta t]du\right) \times \tag{6.60}$$

$$\left(exp\left(-\int_{t_{j+1}}^{s} \sigma(t_{j_1}, u)\sqrt{\Delta t}du\right) + exp\left(\int_{t_{j+1}}^{s} \sigma(t_{j_1}, u)\sqrt{\Delta t}du\right)\right)$$

After taking natural logarithms and differentiating with respect to s[26]

$$\alpha(t_j, s)\Delta t = \frac{\partial}{\partial s} \ln\left(cosh\left(\int_{t_{j+1}}^{s} \sigma(t_j, u)\sqrt{\Delta t}du\right)\right) \tag{6.61}$$

with $cosh(x) = (e^{x} + e^{-x})/2$.

Once the path-dependent interest rate tree is constructed, contingent claims can easily be valued by taking the usual discounted expectation of the claim at maturity. Similar to the models illustrated in the second section of Chapter 6, the interest rate tree has only to span the maturity of the claim, because at each node the whole term structure of forward rates is available. To restrict the number of nodes in the tree to avoid computational problems, it may be necessary to vary the different time intervals.

NOTES

1 To increase the accuracy of this approach by means of, for example, the Control Variate Technique, see Boyle (1977) and Hull and White (1988).

2 For additional restrictions regarding the drift and volatility functions, see Nelson and Ramaswamy (1990, Assumptions 7, 9 and 10).

3 By choosing the binomial probabilities in this way, the instantaneous drift term is matched exactly, but the instantaneous volatility is not. However, the central and noncentral second moment converge to the same limit as $\Delta t \downarrow 0$, because the difference is of the order $\mathcal{O}(\Delta t^2)$.

4 See Nelson and Ramaswamy (1990, Theorem 3).

5 In the models illustrated in the next section, bond values are a known explicit function of the instantaneous interest rate. To value an option on a bond, therefore, an interest rate tree must be constructed that spans the maturity of the option, as the exercise value of the option at maturity can be directly calculated.

6 See Nelson and Ramaswamy (1990, Theorem 4) for conditions regarding convergence in case of stock options. The results established easily translate to European and American call and put options on discount bonds.

7 See n. 3.

8 See Tian (1991, Equation (9)).

9 See Tian (1991, Section 4.3).
10 Note that with three probabilities, the first and second moment of the distribution can be matched exactly in discrete-time. See also n. 3.
11 In addition, Hull and White (1990a, p. 94), suggest taking

$$\frac{\Delta t}{\Delta x^2} = \frac{1}{3}$$

as $\Delta t \downarrow 0$ to increase convergence.

12 See Tian (1991, Section 4.3).
13 See n. 6.
14 See n. 5.
15 See, for example, Geske and Shastri (1985), Nelson and Ramaswamy (1990) Tian (1991), Hull and White (1990a), Boyle (1988), Boyle *et al.* (1989) and Amin (1991).
16 An interesting and necessary extension of the numerical analysis is the determination of the sensitivity of the results obtained with respect to this benchmark value.
17 See also Chapter 5, n. 18 for the general multi-factor case.
18 This different discrete approximation was first discussed in de Munnik (1994a).
19 In Heath *et al.* (1990a, p. 434–5), convergence of the limiting process of the forward rate was obtained by equating the constant martingale probability π to the actual probability q. However, one can show that convergence can be established more generally. In their notation, the discrete martingale probability should equal

$$\pi(j\Delta) = q + \lambda(j\Delta)\sqrt{\Delta}\sqrt{q(1-q)} \quad j = 1, \ldots, \bar{t}$$

where the so-called market price of risk $\lambda(j\Delta), j = 1, \ldots, \bar{T}$ is deterministic and bounded on $[0, T]$.

20 In Amin (1991), the entire forward rate curve is similarly discretely approximated. At maturity of the claim, therefore, the term structure of interest rates is known. However, practical implementation is facilitated in our approximation, as yields are a closed-form solution of the short-rate only, whereas in Amin, yields are a sum of different forward rates.
21 See again Chapter 5, n. 18 with $n=2$ and $\kappa_1=0$.
22 This was first noted by Musiela *et al.* (1990).
23 This method is basically an extension of the approach described in Hull (1989, p. 225–6) for the case of stock options.
24 See, again, Chapter 5, n. 18 with $n=2$ and $\kappa_1=0$ or Heath *et al.* (1992, p. 20).
25 See Heath *et al.* (1990b, p. 75).
26 The volatility function given by equation (6.58), may give some problems as it is not differentiable at $f(t, s)=f_{Max}$. However, this function can be approximated with any desired degree of accuracy by

$$\sigma(t,s,f(t,s)) = \begin{cases} \sigma f(t,s) & \mathit{If} \\ f(t,s) \le f_{Max} - \epsilon & \\ a_1 f(t,s)^2 + a_2 f(t,s) + a_3 & \mathit{If} \\ f_{Max} - \epsilon < f(t,s) < f_{Max} + \epsilon & \\ \sigma f_{Max} & \mathit{If} \\ f(t,s) \ge f_{Max} + \epsilon & \end{cases}$$

with $\epsilon > 0$

$$a_1 = -\frac{1}{4\epsilon}\sigma$$

$$a_2 = \frac{f_{Max} + \epsilon}{2\epsilon}\sigma$$

$$a_3 = \left(-\frac{f_{Max}^2}{4\epsilon} + \frac{f_{Max}}{2} - \frac{\epsilon}{4}\right)\sigma$$

It is readily verified that this approximation is differentiable for all $f(t, s)$.

Part II
EMPIRICAL RESULTS OF THE ESTIMATION OF INTEREST RATE DYNAMICS

7
ESTIMATING THE TERM STRUCTURE OF INTEREST RATES: A TIME SERIES ANALYSIS

In this chapter, the models of Vasicek (1977) and Cox, Ingersoll and Ross (1985) are estimated by means of a time series analysis.[1] Although these two models represent only some specific models within the class of endogenous term structure of interest rates models, there are several reasons for concentrating the empirical analysis on these specific interest rate processes.

The models within the class of the endogenous term structure of interest rate models imply both a stochastic dynamics of spot rates over time and a particular shape of the yield curve at some valuation date. These models, therefore, allow for an interesting comparison between the results of a time series analysis of spot rates and the results of a cross-sectional analysis of bond prices to assess the implications of the shape of the yield curve on the distribution of future spot rates.

The derivation of the different models in this thesis is based principally on the continuous time evolution of interest rates. The obviously discrete time observation of spot rates, therefore, requires a discrete time distribution that is equivalent to its continuous time counterpart. Although one can assume that in the case of daily observations of spot rates the discrete time process is approximately equal to the continuous time process, this chapter shows that an exact aggregation over time seems to be necessary in the case of longer maturity spot rates. This exact aggregation over time to determine the discrete time distribution and its corresponding moments results in closed-form solutions for the models of Vasicek and Cox, Ingersoll and Ross (CIR), facilitating the actual time series analysis.

The main difference between the Vasicek model and the CIR model is the specification of the variance of the spot rates. In the Vasicek model, this variance is constant for a given maturity; because of the resulting normality of interest rates, positive probabilities are assigned to negative rates. In the CIR model, however, the discrete time variance is a linear function of the spot rate with a positive, constant intercept, and, as shown in Chapter 5, this heteroscedasticity excludes negative nominal interest rates. The time series analysis of spot rates and the nested variance specifications, therefore, enable an explicit econometric test of the contribution or significance of the exclusion of negative interest rates.

The first section of this chapter briefly reviews and repeats the discrete time distributions of the spot rate under the Vasicek and CIR models. Next follows a discussion and explanation of the econometric methodology to estimate these time series equations. The next section gives a description of the data used to estimate the different models, while the third section thoroughly discusses the results of the estimation. The last section illustrates the economic implications of these results for the term structure of interest rate volatilities by valuing options on bonds.

ESTIMATION OF THE MODELS

Chapter 5 derived the conditional distribution of the spot rate under the Vasicek model (see p. 69–76). Conditional upon the information at time t, the spot rate $R(t+\Delta t, \tau)$ with maturity τ at time $t+\Delta t$, is normally distributed and follows the following first order autoregressive process

$$R(t+\Delta t,\tau)=b_0+b_1(R(t,\tau)-b_0)+\epsilon(t+\Delta t) \tag{7.1}$$

The parameter b_0 is the unconditional mean of the spot rate, that is:

$$b_0=\lim_{s\to\infty}E_P(R(s,\tau)\mid\mathcal{F}_t)=R(t,\infty)+\frac{1-e^{-\kappa\tau}}{\kappa\tau}\times$$

$$(\theta-R(t,\infty))+\frac{\sigma^2}{4\kappa^3\tau}(1-e^{-\kappa\tau})^2$$

while b_1 reflects the mean reversion of interest rates

$$b_1=e^{-\kappa\Delta t}$$

The disturbances are conditionally normally distributed with mean

$$E_P(\epsilon(t+\Delta t)\mid\mathcal{F}_t)=0$$

and constant variance

$$E_P(\epsilon(t+\Delta t)^2\mid\mathcal{F}_t)=s^2=\left(\frac{1-e^{-\kappa\tau}}{\kappa\tau}\right)^2(1-e^{-2\kappa\Delta t})\frac{\sigma^2}{2\kappa}$$

Not all of the parameters of the Vasicek model can be identified from the linear regression (7.1). The identifiable structural parameters are κ and σ^2, which can be expressed as a function of b_1 and s^2,

$$\kappa=-\frac{\ln(b_1)}{\Delta t} \tag{7.2}$$

$$\sigma^2=2s^2\kappa^3\tau^2(1-e^{-\kappa\tau})^{-2}(1-e^{-2\kappa\Delta t})^{-1} \tag{7.3}$$

Although the unconditional mean b_0 of the spot rate can be estimated, the unconditional mean θ of the instantaneous short-rate and the market price of risk λ are not identifiable separately. The time series estimation, therefore, does not allow for the construction of yield curves and term structures of interest rate volatilities.

The conditional distribution of the spot rate in the case of the CIR model, presented in Chapter 5, also follows a first order autoregressive process:

$$R(t+\Delta t,\tau)=b_0+b_1(R(t,\tau)-b_0)+\epsilon(t+\Delta t) \tag{7.4}$$

The parameter b_0 again represents the unconditional mean of the spot rate,

$$b_0=\lim_{s\to\infty}E_P(R(s,\tau)\mid\mathcal{F}_t)=-\frac{\ln A(\tau)}{\tau}+\frac{B(\tau)}{\tau}\theta$$

while b_1 is the autocorrelation coefficient

$$b_1=e^{-\kappa\Delta t}$$

The disturbances, however, are not conditionally normally distributed. Although the mean is equal to zero:

$$E_P(\epsilon(t+\Delta t)\mid\mathcal{F}_t)=0,$$

the variance is linearly dependent on the lagged spot rate at time t,

$$E_P(\epsilon(t+\Delta t)^2\mid\mathcal{F}_t)=a_0+a_1R(t,\tau)$$

where

$$a_0=\left(\frac{B(\tau)}{\tau}\right)^2\theta\left(\frac{\sigma^2}{2\kappa}\right)(1-e^{-\kappa\Delta t})^2+\left(\frac{B(\tau)}{\tau}\right)^2\left(\frac{\sigma^2}{\kappa}\right)\times$$

$$(e^{-\kappa\Delta t}-e^{-2\kappa\Delta t})\frac{\ln A(\tau)}{B(\tau)}$$

and

$$a_1=\frac{B(\tau)}{\tau}\left(\frac{\sigma^2}{\kappa}\right)(e^{-\kappa\Delta t}-e^{-2\kappa\Delta t})$$

As can readily be seen, the Vasicek model is nested within the CIR model by setting $a_1=0$. In this case, either interest rates are deterministic or the resulting stochastic dynamics can only be captured by the Vasicek model. This hypothesis can be tested with a Lagrange Multiplier Test by computing NR^2 of a regression of the squared residuals of the Vasicek model on a constant and the lagged interest rate, where N denotes the number of observations and R^2 the coefficient of multiple correlation.[2]

The first order autoregressive process of the spot rate (7.4) can be estimated by feasible GLS, which consists of a two-step OLS procedure. In the first step, the autoregressive process is estimated by OLS. The resulting squared residuals are then regressed on a constant and the lagged interest rate to obtain first round estimates of the coefficients a_0 and a_1. Efficient estimates are obtained by running a second round of regressions, where the estimated variances are used to remove the heteroscedasticity by weighted least squares. Finally, the resulting squared residuals of the weighted least squares regression are regressed on a constant and the lagged spot rate to obtain the final estimates[3] of a_0 and a_1.

The estimated four parameters, b_0, b_1, a_0 and a_1, enable the separate identification of the structural parameters κ, σ, θ and λ.

DESCRIPTION OF THE DATA

For the estimation of the time series equations of the Vasicek and CIR models, daily observations of the nominal Amsterdam InterBank Offered

Table 7.1 Summary statistics of AIBOR data

τ	*Mean*	*St. Dev.*	*Minimum (date)*	*Maximum (date)*
1/12	6.32	1.37	3.54 (88/06/24)	9.66 (90/12/27)
	0.00239	0.0649	−0.375 (85/05/08)	0.375 (85/02/04)
2/12	6.35	1.38	3.82 (88/06/24)	9.61 (90/12/27)
	0.00239	0.0595	−0.313 (87/01/07)	0.375 (85/02/07)
3/12	6.37	1.40	4.03 (88/04/14)	9.60 (90/12/27)
	0.00240	0.0620	−0.375 (85/02/18)	0.375 (85/02/07)
6/12	6.43	1.43	4.11 (88/03/08)	9.58 (90/12/18)
	0.00233	0.0613	−0.500 (85/02/18)	0.250 (85/02/01)
12/12	6.50	1.45	4.25 (88/03/10)	9.57 (90/12/18)
	0.00226	0.0612	−0.375 (85/02/18)	0.250 (85/02/04)

This table contains the summary statistics of the AIBOR data. The sample period is 85/01/02 until 90/12/31, which is equivalent to 1551 daily observations. τ denotes the maturity of each interest rate. The first line corresponding to each maturity date contains the summary statistics of the interest rate, which is measured in percentages, while the second line contains the statistics of the first differences of this interest rate. Minimum and Maximum refer to the minimum and maximum interest rate during this period. In parentheses is the date the minimum and maximum occurred.

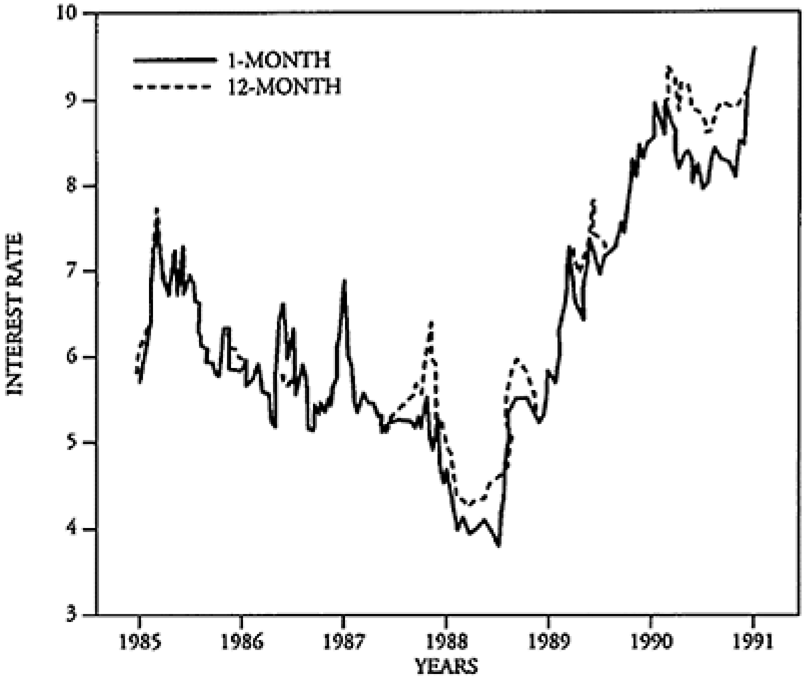

Figure 7.1 The Amsterdam InterBank Offered Rate

This figure shows the one-month and twelve-month Amsterdam InterBank Offered Rate (AIBOR) on a daily basis for the years 1985 through 1991.

Rate (AIBOR) have been used during the years 1985 through 1990. The estimations are carried out for different time series of interest rates with a maturity of one, two, three, six and twelve months, respectively.

Table 7.1 contains summary statistics of these interest rates and the corresponding first difference. From this table and Figure 7.1, it is clear that interest rates do not exhibit a significant upward drift. In addition, the mean interest rate and the corresponding standard deviation during this period is an increasing function of the time-to-maturity. However, the minimum and maximum value of the different interest rates show exactly an opposite relationship.

RESULTS OF THE ESTIMATION

The results of the time series estimation of the Vasicek and CIR models for the different interest rate series are shown in Table 7.2.

The existing literature on time series econometrics makes it clear that a long time series, spanning many years, is needed to estimate the mean reversion parameter, and to test the hypothesis κ=0. Since a discrete

Table 7.2 Results of time series analysis The model of Vasicek

	1/12	*2/12*	*3/12*	*6/12*	*12/12*
$\hat{b}_0$	7.507 (8.210)	7.505 (7.222)	7.449 (6.916)	7.403 (6.246)	7.517 (6.685)
$\hat{b}_1$	0.99799 (0.0030)	0.99794 (0.0029)	0.99778 (0.0031)	0.99762 (0.0031)	0.99778 (0.0029)
$\hat{s}^2 \times 10^{-3}$	4.227 (0.1518)	3.546 (0.1273)	3.854 (0.1384)	3.777 (0.1356)	3.761 (0.1351)
$\hat{\kappa}$	0.5027 (0.2974)	0.5158 (0.2679)	0.5569 (0.2775)	0.6957 (0.2679)	0.5558 (0.2640)
$\hat{\sigma}^2$	1.1040 (0.0491)	0.9673 (0.0567)	1.1080 (0.0893)	1.2660 (0.1859)	1.6010 (0.4669)
	The model of Cox, Ingersoll and Ross				
	1/12	*2/12*	*3/12*	*6/12*	*12/12*
$\hat{b}_0$	7.375 (7.775)	7.337 (6.799)	7.270 (6.480)	7.169 (5.739)	7.265 (6.082)
$\hat{b}_1$	0.99789 (0.001202)	0.99782 (0.001089)	0.99764 (0.001132)	0.99743 (0.001103)	0.99755 (0.001098)
$\hat{a}_0 \times 10^{-3}$	2.313 (1.457)	1.470 (1.125)	1.242 (1.277)	0.304 (1.256)	0.340 (1.103)
$\hat{a}_1 \times 10^{-4}$	3.036 (2.254)	3.278 (1.718)	4.105 (1.958)	5.414 (1.908)	5.271 (1.656)
$\hat{s}^2 \times 10^{-3}$	4.551 (0.5022)	3.873 (0.4196)	4.225 (0.4771)	4.184 (0.5050)	4.169 (0.4926)
LM(1)	0.82	1.50	2.32	5.18*	6.67*
$\hat{\kappa}$	0.5271 (0.3011)	0.5445 (0.2730)	0.5906 (0.2836)	0.6439 (0.2764)	0.6125 (0.2751)
$\hat{\sigma}^2$	0.1546	0.1323	0.1457	0.1463	0.1438

These tables contain the estimated coefficients and structural parameters of the Vasicek and CIR models. The sample period is 85/01/02 until 90/12/31, which is equivalent to 1551 daily observations of AIBOR. The maturity is expressed in years, while the estimated coefficients and structural parameters are measured in percentage points. The standard errors are denoted in parentheses. *LM*(1) denotes the value of the test statistic of the Lagrange Multiplier test with one degree of freedom. An asterix denotes significance at the 5 per cent level (3.84).

time representation of the processes of the risk-free rate is a first order autoregression, a test for κ=0 amounts to testing for a unit root. For the Vasicek model, the test in Fuller (1976) could be used for this purpose. But, with only six years of data, the critical value for κ is 2.35, which is implausibly large.[4] For smaller values of κ, the unit root hypothesis (implying the Merton (1973) model with zero drift for the term structure) cannot be rejected. The mean reversion parameters are approximately the same for all maturities, as should be the case. Neither point estimates nor the standard errors are sensitive to the heteroscedasticity correction applied in estimating the CIR model.

The 1551 observations in the six-year sample are informative, though, about the volatility of interest rates. The asymptotic standard error on $\hat{s}^2$ in the Vasicek model is of the order $\hat{s}^2\sqrt{2/N}$, which is about 3.5 per cent of the estimated value. For the CIR model, the unconditional variance $\hat{s}^2 = \hat{a}_0 + \hat{a}_1\hat{b}_0$ is of the same order of magnitude.

The Lagrange Multiplier Test rejects the constant-volatility Vasicek model in favor of the heteroscedasticity implied by the CIR model only for the two longest maturities. Compared to the results of Chan *et al.* (1991, 1992) for the US Treasury Bill rates and the Japanese Gensaki rate, this result for daily data is surprising. Based on monthly data, the conditional volatility of the one-month T-Bill rate and three-month Gensaki rate appears to be highly sensitive with respect to the level of the lagged interest rate.

The last row in each of the tables presents estimates of the structural variance parameter σ^2. The implied variance of the Vasicek model is estimated precisely for all individual maturities. The estimates, however, differ across maturities. The variance of the twelve-month rate is about 1.5 times the variance of the one-month rate. The estimates of σ_2 are also sensitive to the method of aggregation, as noted in the introduction to this chapter. The approximate estimates $\hat{\sigma}^2 = \hat{s}^2/\Delta t$ are downward-biased by between 5 per cent for the one-month rate and 40 per cent for the twelve-month rate.

For the CIR model, it is impossible to solve the set of non-linear equations for the structural parameters σ^2, θ and λ, given the unrestricted estimates of a_0, a_1, b_0 and b_1, as one of the equations on the system is nearly redundant. The actual estimates in Table 7.2 are obtained by approximating θ (the unconditional mean of the instantaneous short-rate) by $b_{0\tau}$ (the unconditional mean of the yield on a discount bond with maturity τ).[5] The error of this approximation is likely to be very small given that the term structure was almost flat over most of the sample period according to the results derived in the next chapter.

Although the econometric analysis of this chapter focuses only on two specific models within the class of endogenous term structure of interest rates models and no comparison has been made between these models and other specifications in which the conditional variance is dependent on higher order lagged interest rates, the results are encouraging. Given the rejection of the heteroscedasticity model of Cox, Ingersoll and Ross in the case of short maturity interest rates and the robustness of conditional variance specifications with respect to structural breaks in monetary policy, as shown by Chan *et al.* (1992, pp. 13–14), the primary objective of valuing interest rate derivative securities and of comparing the above results to a cross-sectional analysis of the same term structure models seems to be reasonably supported.

OPTION PRICING

The main result of comparing the time series analysis of the Vasicek model and CIR model is the statistical rejection of the constant volatility for longer maturity interest rates. Despite the previous research of Chan *et al.* (1991, 1992) concerning the US Treasury Bill rate and the Japanese Gensaki rate resulting in opposite conclusions, the heteroscedasticity implied by the CIR model is shown to be insignificant in comparison to the Vasicek model for maturities up to three months.

An interesting extension to this comparison is the implication of the results of the estimations of the different models for the values of call options on bonds. As options are critically dependent on the term structure of volatilities, a comparison of the two models with respect to these option prices based on the estimations for short maturity interest rates, provides some insight into the implications for longer maturity volatilities.

The actual calculation of the options based on the estimations discussed in the previous sections requires some additional assumptions. In the case of the Vasicek model, the value of an option on a discount bond depends on the parameters κ, σ^2 and the infinite maturity yield $R(t, \infty)$. Because the estimated coefficients allow only for a direct calculation of the first two parameters, the infinity maturity yield can only be determined by making an additional assumption. For this purpose, the one-month and two-month AIBOR are used to calculate the infinite maturity yield on a daily basis using equation (5.21). The instantaneous spot rate on a daily basis is then obtained using the one-month AIBOR on a daily basis and the average infinite maturity yield over the sample period. In the case of the CIR model, the separate identification of the parameters κ, σ^2 is only possible by assuming a flat term structure of interest rates because the system of equations, of which these parameters are a solution, is nearly redundant. To determine the infinite maturity yield $R(t, \infty)$, it is assumed that the market price of risk is zero for the same reason. The instantaneous spot rate on a daily basis is then obtained by using equation (5.34) and again the one-month AIBOR on a daily basis and the infinite maturity yield.[6]

Figure 7.2 shows the CIR European call option values as a function of the corresponding Vasicek option values. To be able to compare the time series option values with the cross-sectional option values of the next chapter and to concentrate as much as possible on the time value of the option, the maturities of the option and the underlying bond are four and eight years, respectively. The exercise price of the option is equal to the forward price of the underlying value, while the face value is 100.

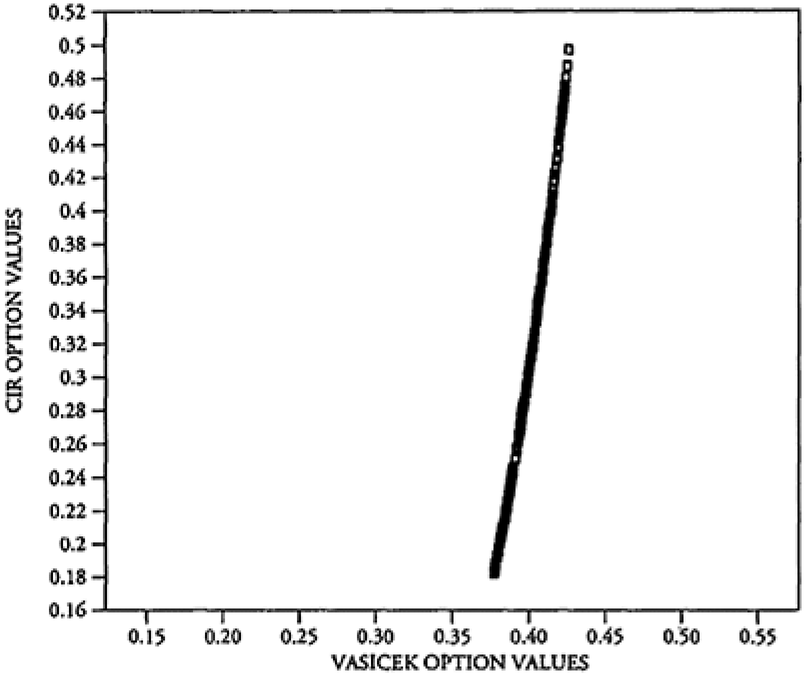

Figure 7.2 Time series option values

This figure shows the European at-the-money call option values on a daily basis for the years 1985 through 1991 (1551 option values) based on the results of the time series estimation for the Vasicek and CIR models. The values of the options for the latter model are calculated numerically using the method of Nelson and Ramaswamy, described in Chapter 6. The maturity of the underlying bond is eight years, while the maturity of the option is four years.

Although the determination of the option values is not justified theoretically because of the additional assumptions, it is interesting to note the relatively low option values present for both models. Although the values of the underlying long term bond are slightly higher and more sensitive with respect to changes in the one-month AIBOR in the case of the CIR model, option values are generally higher in the case of the Vasicek model.

NOTES

1 This chapter is based on "Cross-sectional versus Time Series Estimation of Term Structure Models: Empirical Results for the Dutch Bond Market", J.F.J.de Munnik and P.C.Schotman, 1994, Journal of Banking and Finance, Vol. 18, No. 5, pp. 997–1025.

2 For a thorough discussion of the Lagrange Multiplier Test, see Harvey (1990, pp. 172–3).

3 For a discussion of this two-step procedure and the corresponding efficiency of the estimates, see Maddala (1988, pp. 170–1).

4 According to Table 8.5.1 in Fuller (1976, p. 371), the 5 per cent critical value of the test statistic $N(\hat{\rho}-1)$ is −14.1, where *N* is the number of observations and $\hat{\rho}$ the estimated autocorrelation. Substituting

$$\kappa = -\frac{\ln(\rho)}{\Delta t}$$

using ln(1+*x*) ≈ *x* for small *x*, and noting that *N*Δ*t* is the number of years, we find the inequality

$$\kappa \geq 2.35$$

5 Equating $\hat{\theta}$ to $\hat{b}_{0\tau}$ implies

$$\frac{\ln A(\tau)}{\tau} = 0 \quad \text{and} \quad \frac{B(\tau)}{\tau} = 1$$

The estimated value of the unconditional variance of the yield then gives

$$\hat{\sigma}^2 = 2\frac{\hat{s}^2}{\hat{b}_{0\tau}} \frac{\hat{\kappa}}{1-e^{-2\hat{\kappa}\Delta t}}$$

6 To assess the impact of longer maturity AIBOR rates on the results, option prices were also calculated using these different rates. The results, however, hardly changed.

8
ESTIMATING THE TERM STRUCTURE OF INTEREST RATES: A CROSS-SECTIONAL ANALYSIS

In the previous chapter, two models within the class of endogenous term structure of interest rate models were estimated by means of a time series analysis. Based on the Amsterdam InterBank Offered Rate for different maturities, the constant variance of interest rates specification of the Vasicek (1977) model was not rejected in favor of the heteroscedastic variance specification implied by the model of Cox, Ingersoll and Ross (CIR) (1985) for maturities of up to three months. Given the estimations of the structural parameters of these models, option prices were also calculated to assess the impact of the results of the time series analysis on long-maturity volatilities. Although this latter analysis is not theoretically completely justified and some additional assumptions have to be made, it turned out that, over the sample period, option values based on the CIR model are more sensitive to the value of the instantaneous spot rate than are those option values based on the Vasicek model.

As already mentioned, there are several reasons for concentrating this empirical analysis on the specific interest rate models of Vasicek and Cox, Ingersoll and Ross. The stochastic differential equation of the spot rate and the term structure of interest rates implied by these dynamics and the well-known no-arbitrage conditions, allow for an interesting comparison between a time series analysis of spot rates and the stochastic behavior of interest rates implied by a particular shape of the yield curve. In addition, the two models considered differ with respect to their variance specification of spot rates; it is interesting, thus, to investigate empirically the increased complexity resulting from the exclusion of negative nominal interest rates in the case of the CIR model, as opposed to the normally distributed nominal interest rates in the case of the Vasicek model.

This chapter compares the two above-mentioned models empirically by means of a cross-sectional analysis for the Dutch Government Bond Market.[1] The first section briefly reviews the functional specifications of the implied yield curve and discusses the estimation technique, and is followed by a description of the data in the second section. The third section discusses and illustrates the results of the estimation. In the fourth section the implications of these results for the implied term structure of interest rate volatilities are discussed by means of calculating the values of options on discount bonds. In the final section, a particular week of the sample period is analyzed in detail to highlight some problems and estimation difficulties.

ESTIMATION OF THE MODELS

In the Vasicek model, discussed in Chapter 5, the spot rate $R(t, \tau)$ at time t with maturity τ can be written as a weighted average of the instantaneous short-term rate of interest and the infinite maturity yield plus a linear function of interest rate volatility, that is,

$$R(t,\tau) = w_1(\kappa,\tau)r(t) + (1 - w_1(\kappa,\tau))R(t,\infty) + w_2(\kappa,\tau)\sigma^2 \tag{8.1}$$

with

$$w_1(\kappa,\tau) = \frac{1 - e^{-\kappa\tau}}{\kappa\tau}$$

$$w_2(\kappa,\tau) = \frac{(1 - e^{-\kappa\tau})^2}{4\kappa^3\tau}$$

As can easily be shown, the two weighting functions satisfy the following conditions:

$$\lim_{\tau\to 0} w_1(\kappa,\tau) = 1 \quad \lim_{\tau\to\infty} w_1(\kappa,\tau) = 0 \quad 0 \leq w_1(\kappa,\tau) \leq 1$$

$$\lim_{\tau\to 0} w_2(\kappa,\tau) = 0 \quad \lim_{\tau\to\infty} w_2(\kappa,\tau) = 0 \quad w_2(\kappa,\tau) \geq 0$$

At both ends of the maturity spectrum, the effect of volatility on the shape of the yield curve is zero. In between these two points, volatility is increasing the level of interest rates, causing a curvature of the yield curve. For reasonable values of the parameters κ and σ^2, however, the contribution to this curvature of the volatility is negligible. If, for example, $\kappa=1$ and $\sigma^2=1$, which is reasonable given the results of the time series analysis of Chapter 7,[2] the maximum contribution of volatility to the level of interest rates is less than one basis point. To have some impact on the yield curve, it takes extreme values for either the mean reversion parameter κ or the volatility σ.

Because only coupon-paying bonds are traded in the Dutch Government Bond Market, the values of these bonds have to be expressed in terms of zero-coupon bonds. Consider, therefore, a coupon bond at time t, $P(t, \tau, c)$, which entitles the holder to a vector of n cash flows $c=(c_1,\dots, c_n)^T$ with corresponding payment dates $\tau=(\tau_1,\dots, \tau_n)^T$. The value of such a bond or dirty price in terms of the different discount bonds is

$$P(t, \boldsymbol{\tau}, \mathbf{c}) = \sum_{j=1}^{j=n} c_j P(r(t), t, \tau_j) \tag{8.2}$$

To estimate the parameters of the Vasicek model at time t, it will be assumed, similar to Brown and Dybvig (1986), that the quoted bond price $P^*(t, \tau, c)$ deviates from the model price $P(t, \tau, c)$ by a zero-mean error $\epsilon_{(c,\tau)}(t)$A justification of this stochastic error term is the presence of a measurement error due to the bid-ask spread. Assume that the errors are i.i.d., allowing cross-sectional estimation of equation (8.2) by NLS.[3] Since the instantaneous spot rate *r(t)* is unobservable, *r(t)* is treated as an unknown parameter, which is estimated jointly with $(\tilde{\theta}, \sigma^2, \kappa)$, as in Brown and Dybvig. The market price of risk λ. is not individually observable.

In the case of the CIR model, the spot rate can also be expressed as a weighted combination of the instantaneous spot rate and the infinite maturity yield:

$$R(t,\tau) = w_1(\tilde{\kappa}, \sigma, \tau) r(t) + w_2(\tilde{\kappa}, \sigma, \tau) R(t, \infty) \tag{8.3}$$

$$w_1(\tilde{\kappa}, \sigma, \tau) = \frac{B(\tau)}{\tau}$$

$$w_2(\tilde{\kappa}, \sigma, \tau) = -\frac{(\tilde{\kappa} + \gamma)}{2\kappa\theta} \frac{\ln A(\tau)}{\tau}$$

The two weighting functions in this case obey

$$\lim_{\tau \to 0} w_1(\tilde{\kappa}, \sigma, \tau) = 1 \quad \lim_{\tau \to \infty} w_1(\tilde{\kappa}, \sigma, \tau) = 0 \quad w_1(\tilde{\kappa}, \sigma, \tau) \geq 0$$

$$\lim_{\tau \to 0} w_2(\tilde{\kappa}, \sigma, \tau) = 0 \quad \lim_{\tau \to \infty} w_2(\tilde{\kappa}, \sigma, \tau) = 1 \quad w_2(\tilde{\kappa}, \sigma, \tau) \geq 0$$

Contrary to the findings of the Vasicek model, the CIR spot rate is not a weighted average of the instantaneous short and infinite maturity yield. The two different weighting functions are also a function of the volatility of the spot rate. Although not directly clear, it can be shown that volatility is causing a curvature of the yield curve in this case, too.

Similar to the estimation of the Vasicek model, the instantaneous spot rate *r(t)* can be treated as an unknown parameter and equation (8.2) can be estimated by NLS. From this cross-sectional estimation, the parameters $(\tilde{\kappa}, \sigma^2, \theta\kappa)$can be identified separately. Estimation of the individual parameters of the risk-neutral process of the short-rate using bond prices, however, results in the impossibility of an individual identification of the parameters λ, κ and θ.

DESCRIPTION OF THE DATA

For cross-sectional estimation of the Vasicek and CIR models, data of actively traded Dutch Government Bonds with a remaining maturity of longer than six months for each trading day during 1989 and 1990 has been used. For each of the bonds in the sample, data has been gathered on the clean closing price and accrued interest. This data allowed for the computation of the corresponding cash flow patterns consisting of coupon payments, final repayment and corresponding payment dates.

Table 8.1 provides summary statistics of the data set. The total number of trading days is 507, while the number of actively traded bonds in a day varies between 33 and 47. The longest maturity is about ten years, while the duration of the longest bond is, on average, somewhat more than seven years during the sample period. During most of the period, the yield curve has been flat, and, on average, it has been inverted.[4]

In Figure 8.1, the yield on the longest maturity bond is shown together with the one-month AIBOR during the same period. As already noted above, during most of the sample period, short-term yields have been higher than long-term yields.

Table 8.1 Summary statistics of bond data

	Mean	*St. Dev.*	*Minimum (date)*	*Maximum (date)*
Max. yield	7.99	0.90	6.40 (89/01/02)	9.30 (90/12/31)
Min. yield	7.65	0.90	5.52 (89/01/05)	8.79 (90/09/28)
Sd. yield	0.08	0.04	0.03 (89/03/13)	0.19 (89/12/18)
Spread	−0.14	0.25	−0.95 (89/10/10)	0.67 (89/01/05)
Maturity	9.97	0.06	9.81 (89/10/25)	10.13 (90/03/22)
Duration	7.21	0.25	6.81 (89/10/08)	7.68 (89/01/02)
Number of bonds	40.13	4.12	33 (89/01/02)	47 (90/10/31)

This table contains the summary statistics of the bond data. The sample period is 89/01/02 until 90/12/31, which is equivalent to 507 daily observations. Max. yield, Min. yield and Sd. yield refer to the maximum yield, the minimum yield and the standard deviation of yields on a given day, respectively. Spread is the difference in yield between the bond with the longest and shortest duration on a day. Maturity and Duration give the maximum maturity and duration on a day. In parentheses is the date the minimum and maximum occurred.

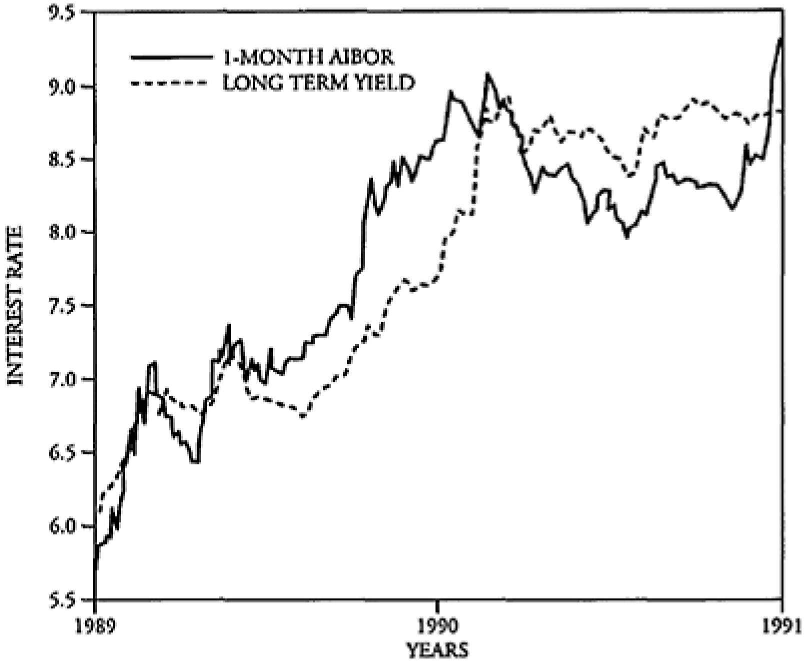

Figure 8.1 Short-term and long-term yields

This figure shows the one-month Amsterdam InterBank Offered Rate and the yield-to-maturity on the longest maturity bond on a daily basis for the years 1989 and 1990.

RESULTS OF THE ESTIMATION

As in Brown and Dybvig (1986), equation (8.2) can be estimated for each trading day of the sample period. Preliminary estimation, however, revealed that the parameters of the CIR model were hardly estimable using data for a single trading day. It was therefore decided to pool the data for the five trading days of the week and assume that the parameter vector was constant over the week. The risk-free rate is allowed to take on a different value each day. For the CIR and Vasicek models, this leaves eight parameters to be estimated each week.

The results of this cross-sectional estimation are summarized in Table 8.2. Both models provide a good fit for bond prices. The average error of both models is 0.18 guilders (par bonds are normalized to 100 guilders). The Vasicek model fits marginally even better than the CIR model does. In Figure 8.2, it is shown that for most weeks the two models fit almost identically, while the fit is almost fairly constant for all weeks.

Because of the frequent occurence of extreme outliers, the mean and standard deviation of the parameter estimates over the 104 weeks are not

Table 8.2 Results of cross-sectional analysis The model of Vasicek

	$\hat{r}(t)$	$\hat{\kappa}$	$\hat{\hat{\theta}}$	$\hat{R}_{\infty}$	$\hat{\sigma}^2$	$\hat{\sigma}$	*s.e.*
Mean	8.28	1.23	9.27	6.45	2633.65	13.07	0.1808
St. Dev.	2.55	1.81	3.96	3.42	12228.08	49.63	0.0366
Minimum	0.60	0.01	2.49	0.50	0.05	0.22	0.1017
Maximum	15.78	5.00	31.49	21.46	83392	288.78	0.3046
Median	8.73	0.19	7.77	6.97	0.05	0.22	0.1726
Range	2.16	1.57	2.59	2.46	0	0	0.0468
The model of Cox, Ingersoll and Ross							
	$\hat{r}(t)$	$\hat{\hat{\kappa}}$	$\hat{\theta}$	$\hat{R}_{\infty}$	$\hat{\sigma}^2$	$\hat{\sigma}\sqrt{\hat{r}(t)}$	*s.e.*
Mean	8.68	0.84	9.63	6.63	292.17	17.23	0.1824
St. Dev.	2.45	1.40	4.80	2.56	1371.0	70.35	0.0376
Minimum	5.58	0.005	1.87	0.85	0.01	0.08	0.1017
Maximum	22.01	5.00	27.98	17.57	8862.30	431.63	0.3047
Median	8.76	0.15	7.90	6.90	0.01	0.29	0.1734
Range	2.21	1.16	2.87	2.01	0	0.05	0.0466

This table contains the cross-sectionally estimated structural parameters of the models of Vasicek and Cox, Ingersoll and Ross. The sample period is 89/01/02 through 90/12/31, which is equivalent to a total of 104 weekly estimations. Minimum and Maximum refer to the minimum and maximum value of the 104 weeks. Median and Range refer to the median and interquartile range of the 104 estimated values. s.e. refers to the pricing error.

very informative about the typical parameter estimates. The median and interquartile range are more robust measures of location and dispersion. In Figure 8.3, the estimated risk-free rate is almost equal for the two models, except for some severe outliers. The Vasicek model has a number of exceptionally low estimates, while the CIR model leads to some high estimates of *r(t)*. The outliers are highly correlated across parameters. Whenever an outlier occurs for one structural parameter, an outlier for some of the other parameters is always present. The estimated risk-free rate is nearly always above the observed one-month AIBOR.

In most weeks, the other end of the yield curve or the infinite maturity yield can also be estimated reasonably well. For the Vasicek model, there are four upward outliers (the first four weeks of the sample, see Figure 8.3) and a series of downward outliers at the last quarter of 1990, where $R(t, \infty)$ reaches its numerical lower bound of 5 per cent.[5] Again, the estimates of the Vasicek and CIR models are generally quite similar, although in the CIR model $R(t, \infty)$ never approaches negative values, as it does in the Vasicek model.

The mean reversion parameter κ (or $\tilde{\kappa} = \kappa - \lambda\sigma$ in the CIR model) and the volatility parameter σ^2 behave erratically. They usually differ greatly

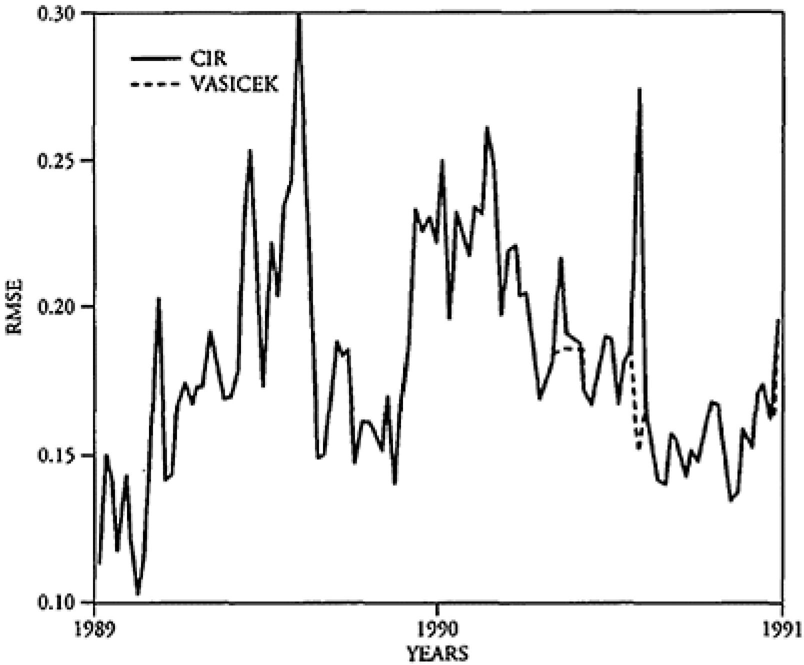

Figure 8.2 Pricing errors

This figure shows the root-mean-square-error on a weekly basis under the models of Vasicek and Cox, Ingersoll and Ross for the years 1989 and 1990.

from week to week with large standard errors.[6] Since the term structure models are valid only for $\kappa > 0$, and since it is numerically impossible to compute bond prices for very small κ, κ had to be restricted to $\kappa > 0.01$ in the Vasicek model and to $\tilde{\kappa} > 0.005$ in the CIR model. This corner solution frequently occurs for both models. The outliers for $R(t, \infty)$ occur only when k is small. In this case, $R(t, \infty)$ is almost unidentified in both models. Too large values of κ are also unacceptable. For $\kappa > 5$ the Hessian of the likelihood function becomes numerically singular. When the optimum of κ falls within the admissible range, the estimates are remarkably similar for the Vasicek and CIR models.

In many cases, σ^2 falls to zero in unconstrained non-linear estimation. For numerical stability, a lower bound of $\sigma^2 > 0.01$ was set for the CIR model and $\sigma^2 > 0.05$ for the Vasicek model. The lower bounds are attained in more than half of the sample weeks. The effect of the lower bound was checked on the fit of the models, which turns out to be negligible. The standard deviation of the residuals never changes by more than one-tenth of a cent if σ^2 is allowed to take on values below the bounds. For most cases the lower bound could also have been set at 0.01

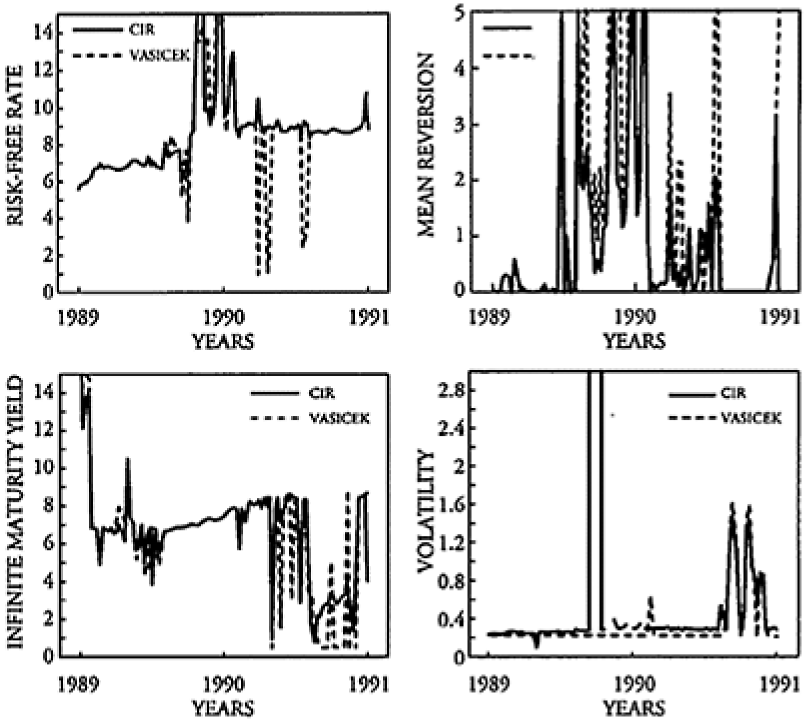

Figure 8.3 Estimated structural parameters

These figures show the estimated values of the different structural parameters on a weekly basis in the case of the Vasicek and CIR models for the years 1989 and 1990.

without significantly deteriorating the fit of the model, indicating that the likelihood function is exceptionally flat.[7] If κ hits its upper bound, the optimization algorithm cannot get any estimate of σ^2, except an extreme outlier ($\sigma^2 > 10{,}000$). For κ at its lower bound, σ^2 always falls to zero. In both cases it seems that the yield curve is already reasonably well described as a weighted average of the instantaneous short-rate *r(t)* and the infinite maturity yield $R(t, \infty)$. Given the discussion in the first section of the effect of volatility on the shape of the yield curve and the observed flat yield curve during the sample period, it is not surprising that the curvature effect implied by the volatility is not present in the data. For those weeks where an unconstrained optimum is not found (mostly in the fourth quarter of 1990), the CIR and Vasicek models are indistinguishable.

The overall impression from the cross-sectional estimates is that the model is overspecified. Without any loss in fit one of the structural parameters can be set at some "reasonable" value and can be optimized over the other. The symptoms of the overparameterization are the frequent occurence of outliers and the near singularity of the Hessian.

In addition, the results do not favor one of the two term structure models. If the parameters are estimated unrestrictedly and revised every week, both models do equally well (or poorly). More general models with more state variables will even lead to more estimation difficulties, since these simple one-factor models are already too flexible. To discriminate between the models, more restrictions are needed; for instance, requiring the structural parameters to be constant over a much larger time period than one week, such as is the case in Brown and Schaefer (1994).

Given the results of the cross-sectional estimation, a third general conclusion is that the implied process of the risk-free rate is close to a random walk without drift and almost deterministic. This conclusion might be specific for the data set, which contains many days with flat term structures. The low implied volatility conflicts strongly with the time series estimates. Interest rates have been quite volatile during the sample period, but the yield curve has usually shifted up and down with movements in the short-rate. In comparison to similar estimations of the CIR model in the case of US Treasuries and British Government Index-Linked bonds, carried out by Brown and Dybvig (1986) and Brown and Schaefer (1994), respectively, this result is remarkable. Based on US Treasury issues for the period 1952 through 1983, cross-sectional estimation of the volatility of the CIR model on a daily basis seems to correspond quite well to the volatility resulting from a time series estimation. Similar to our results, however, the implied values of the short-term spot rate are generally higher than the observed values. Using real interest rates derived from British Government Index-Linked bonds for the period 1981 through 1989, Brown and Schaefer show the infinite maturity yield to be fairly stable. The unconstrained estimated value of the mean reversion parameter also behaves erratically during their sample period and often becomes negative.

OPTION PRICING

Although the Vasicek and CIR models differ principally in their specification of the stochastic differential equation of the short-rate and the resulting value of a discount bond as a function of the structural parameters, both models provide a good fit of the term structure of interest rates for the Dutch Government Bond Market.

As was the case in Chapter 7, it is interesting to investigate and compare the implications of the estimation of the yield for the term structure of interest rate volatilities by valuing European call options on discount bonds. Because the cross-sectional analysis of the term structure of interest rates yields estimations of the parameter values of the stochastic process of the short-rate in a risk-neutral economy, no additional assumptions have to be made in order to calculate the option prices, as in Chapter 7.

In Figure 8.4, the values of a European at-the-money call option on a discount bond under the CIR model as a function of the corresponding option values under the Vasicek model are shown. To cover the relevant maturity spectrum of the different bonds in the sample and to concentrate on the time value of the option, the maturities of the option and the underlying bond are four and eight years, respectively. In addition, the exercise price of the option is equal to the forward price of the underlying bond. In case option values cannot be computed because of numerical problems resulting from extreme structural parameter values, these weeks from the sample are simply ignored.

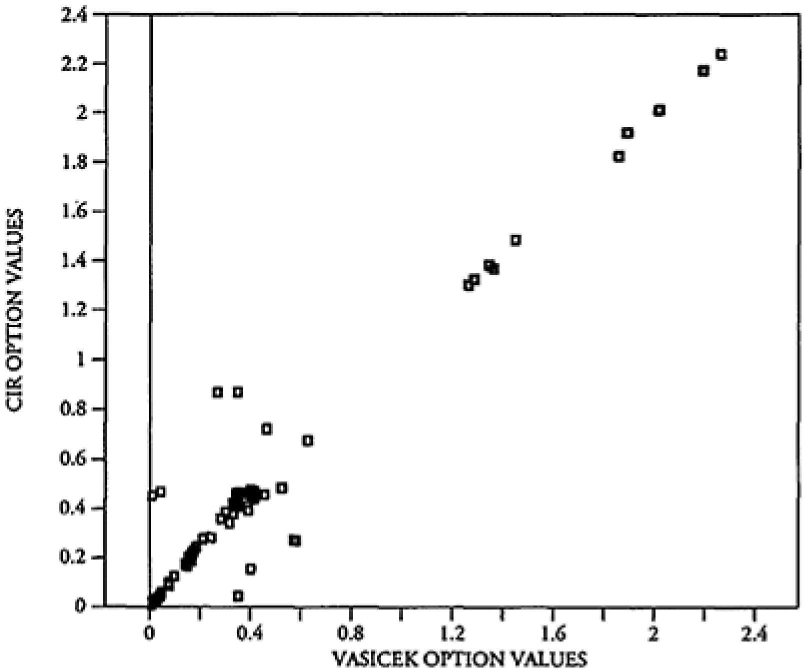

Figure 8.4 Cross-sectional option values

This figure shows the European at-the-money call option values on a daily basis for the years 1989 and 1990, or 403 option values, based on the results of the cross-sectional estimation in the case of the Vasicek and CIR models. If possible, the values of the options in the latter model are calculated numerically using the method of Nelson and Ramaswamy, described in Chapter 6. Otherwise, these days are removed from the sample. The maturity of the underlying bond is eight years, while the maturity of the option is four years.

The general conclusion from this figure is the strong one-to-one correspondence of the implications of the estimation for the valuation of options between both models. Except for a few outliers, the CIR option values are almost equal to the corresponding Vasicek option values, suggesting a significant equivalence between both term structures of interest rate volatilities.

A CLOSER LOOK AT A SELECTED WEEK

The cross-sectional analysis of the term structure of interest rates on a weekly basis often results in estimation problems because of an overspecification of both models. As already noted in the discussion of the results, pooling of different days for longer periods of time such as quarters, years or even the total sample period may reduce these difficulties. To assess the impact of the over-identification more closely and to investigate the properties of the models in more detail, one week has been selected from the sample.

The week 22–6 October 1990 is in many respects typical for the Dutch bond market in our two-year period. The yield curve is flat, and the parameters are very poorly estimated for both models. The best solution is obtained when the mean reversion parameter κ is at its lower bound, although the optimization routine failed to converge for many starting values. In Figure 8.5, the actual and fitted yields of the different bonds in the sample are shown. In addition, the pricing errors and actual and fitted prices for this particular week are presented. The models fit perfectly in the price dimension with a coefficient of multiple correlation R^2 = 0.999 for both models. It fits poorly, however, in the yield dimension, with R^2 = 0.030, implying that the models can explain only 3 per cent of the cross-sectional variation in yields.

To investigate the problems, a grid search was performed over the mean reversion parameter κ. Conditional on a range of values of κ, the other parameters of the model were estimated. Figure 8.6 shows the fit of the model, measured by the root-mean-square-error (RMSE) of the pricing residuals. The CIR and Vasicek models are almost indistinguishable for all values of κ. For both models, the global minimum is obtained for κ at its lower bound. More disturbing is the decrease in the function value for large κ. It explains why it is so hard to obtain point estimates. If the starting value of κ in an optimization algorithm based on derivatives is too large, say larger than 0.3, the algorithm will not converge, and will produce implausibly large estimates of κ, coupled with very small values of σ^2.

The two models are identical not only in fit, but also with respect to other structural parameters. In Figure 8.6, the implied volatility of both models is shown. For the Vasicek model, this is the estimated σ corresponding to each value of κ; for the CIR model, it is $\sigma\sqrt{\bar{r}(t)}$ with $\bar{r}(t)$

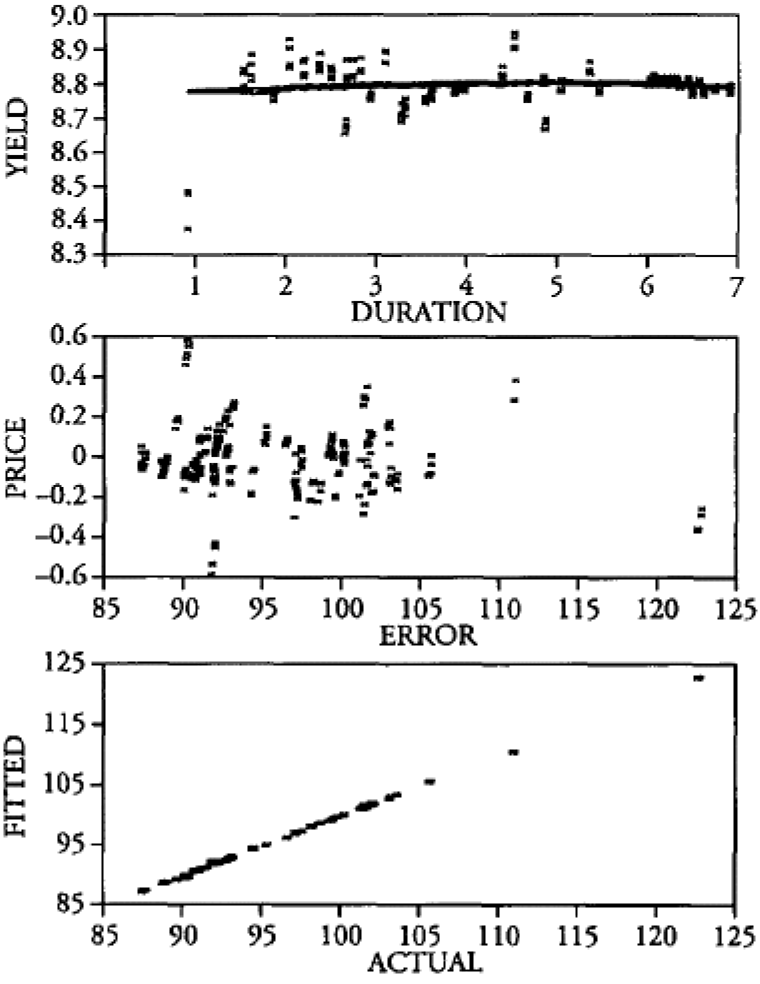

Figure 8.5 Estimation results for a particular week

These figures show some results of the estimation in case of the particular week 22–6 October 1990, which is in many respects typical for the Dutch Bond Market.

the average of the estimated risk-free rates over the five days of the week. In one respect, the week 22–6 October 1990 is special, because there is a range of κ for which the volatility is not at its lower bound.

Finally, Figure 8.6 shows that even the implied option values are almost equal for the two models for all κ. The last two results are surprising, since both the cross-sectional yield curves, and the option formulas, are completely different between the two models. Given that the option values are much easier to compute using the Vasicek model, it seems there is little empirical reason to prefer the theoretical advantages of the CIR model over the Vasicek model.[8]

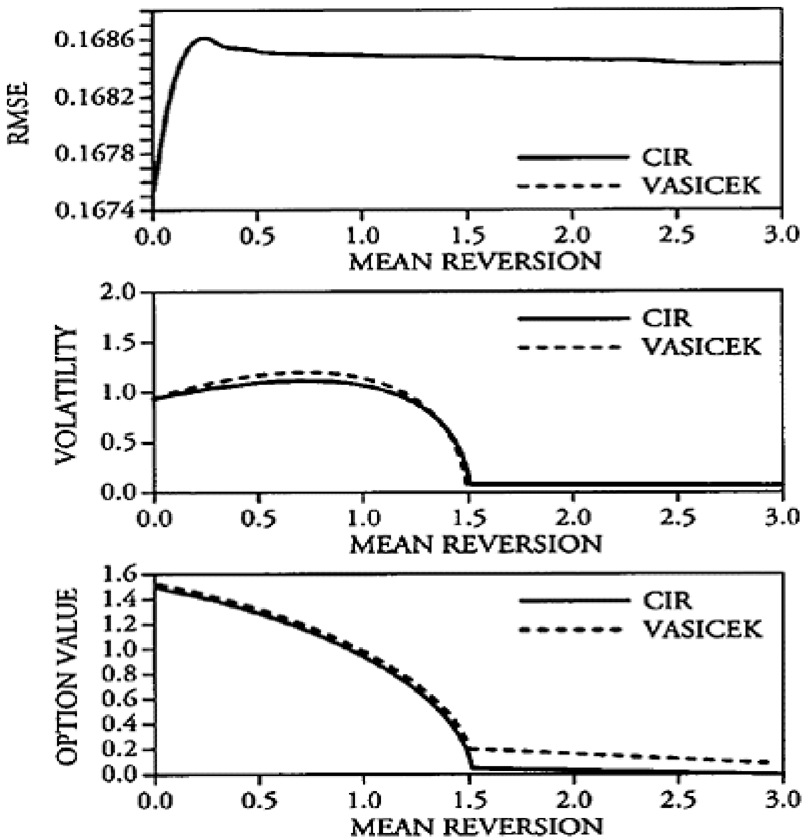

Figure 8.6 Results of grid search over mean reversion

These figures show some results of the grid search over the mean reversion parameter under the Vasicek and CIR models for the particular week 22–6 October 1990, which is in many respects typical for the Dutch Bond Market.

NOTES

1 Like the previous chapter, this chapter is based on “Cross-Sectional versus Time Series Estimation of Term Structure Models: Empirical Results for the Dutch Bond Market”, J.F.J.de Munnik and P.C.Schotman, 1994, Journal of Banking and Finance, Vo1. 18, No. 5, pp. 997–1025.

2 In terms of structural parameters measured in percentages, this accounts for $\kappa=1$ and $\sigma=10^{-4}$.

3 Alternative stochastic specifications are possible. Homoscedastic errors for prices imply heteroscedastic errors for yields. An error in a bond price has the largest effect on yields for short-term bonds. With our specification, long-term yields fit closer than short-term yields. If we aim to fit yields, we would have to assume that the errors in our specification are related to squared duration. This considerably affects estimates of the risk-free rates. Our assumption of homoscedastic errors in prices is consistent with the interpretation of measurement error due to the bid-ask spread, which is constant across maturities.

4 Unfortunately, during the subperiod 1 February 1989, to 3 March 1990, the quoted clean closing price and accrued interest of all bonds in the sample were incorrectly assigned to the next day. For example, the quoted bond price and accrued interest of a bond on 1 February 1989, was assigned to 2 February 1989. Although the pooling of bonds for the five trading days of a week, as discussed and explained in the next section, is therefore incorrect because of a possible “weekend-effect”, a preliminary analysis of the corrected data resulted in negligible changes; because of this, we will not pursue the issue any further.

5 The unconstrained optimization did not converge and led to negative values of $R(t, \infty)$.

6 Standard errors of the parameters were computed from the inverse of the Hessian of the least squares function. The standard error of the risk-free rate is of the order 0.1 per cent, and that of $R(t, \infty)$ is either of the same order or extremely large. The other parameters are never significant. Still, these standard errors probably underestimate the true standard errors for two reasons. Due to the pooling of the five days of the week, it is likely that the errors are autocorrelated. Second, for most weeks a corner solution is obtained where some of the parameters are constrained. In that case, only standard errors conditional on the constrained parameters can be obtained, and these are usually much smaller than the unconditional standard errors.

7 Brown and Dybvig (1986) and Brown and Schaefer (1994) encounter similar problems. They also found a flat likelihood function, while the (indirect) estimates of σ^2 are often negative in the case of Brown and Dybvig.

8 The robustness of this result to different shapes of the yield curve must, of course, be investigated.

9
ESTIMATING THE TERM STRUCTURE OF INTEREST RATE VOLATILITIES: PRINCIPAL COMPONENTS

The use of principal components is a commonly applied statistical technique in the finance literature to determine a reduced number of factors or sources of uncertainty to describe the stochastic movement of the term structure of interest rates over time. Using weekly observations from 1984 until 1988 in the case of US Treasuries, for example, Litterman and Scheinkman (1988), conclude that, based on principal component analysis, a three-factor model explains at least 98 per cent of the variability of excess returns of any zero coupon bond. The first factor essentially represents a parallel change in yields, while the second and third factors represent a change in the steepness and curvature of the yield curve, respectively. These findings have been confirmed by several authors, for example, Garbade (1986), Dybvig (1989) and Heath *et al.* (1990b).

The estimation of the endogenous term structure of interest rate models by means of a time series or cross-sectional analysis, as conducted in the previous two chapters, directly implies a particular shape of the endogenous term structure of interest rate volatilities. As options are critically dependent on both the yield and volatility curves, implied estimation of only the yield curve may ignore or misspecify some of the volatility characteristics mentioned above.

The exogenous term structure of interest rate models avoids these possible problems by allowing for a separate estimation of the yield curve and the term structure of interest rate volatilities. Given a yield curve at some initial valuation date, the endogenous term structure of interest rate volatilities enables the specification or estimation of a volatility curve as a function of a few structural parameters without affecting the initial term structure of interest rates. In the case of the exogenous term structure of interest rate models, an exact matching of the observed or estimated volatility curve can be obtained; this approach enables the application and integration of principal components analysis to the valuation of interest rate derivative securities.

An interesting problem arising from the application of principal component analysis is whether the reduced number of factors used to describe the variability of the term structure over time is sufficient to model the specific volatility structure necessary for the valuation of options on bonds. Although a reduced number of factors fairly accurately describes the general variability of the yield curve over time, this chapter shows that they are generally not sufficient to describe the variability of specific maturity segments upon which options are crucially dependent.[1] As a result, the application of principal components is questionable in determining a relatively small number of factors to describe the movement of the term structure over time and to value interest rate dependent securities, based on this reduced set of principal components.

The first section of this chapter derives the analytical relationship between a principal component analysis and the valuation of European call options on discount bonds under the Jamshidian (1989) model, discussed in Chapter 5. Due to the possibility of an explicit formulation of the distribution of different spot rates, the problems related to the volatility parameter of the option and the reduced number of principal factors can be identified exactly. In the second section, a simple numerical example is presented, showing the dependence of the call option of the different factors to assess the estimation problems for the model of Heath *et al.* (1990b), also presented in Chapter 5. Because an explicit formulation of the distribution of forward rates is not possible and a closed-form solution for the value of an option on a discount bond does not exist, this numerical counter example is the only way of investigating these statistical problems.

THE JAMSHIDIAN MODEL

In the N-factor "variable-volatility" model of Jamshidian (1989) discussed in Chapter 5,[2] the stochastic process of the instantaneous spot rate $r(t)$ at time t [0, T], given the information at time 0 is equal to

$$r(t) = f(0, t) + \sum_{i=1}^{i=N} \delta_i(t) + \sum_{i=1}^{i=N} \sigma_i(t) W_i(t) \qquad (9.1)$$

with

$$\delta_i(t) = \sigma_i(t) \int_0^t \int_u^t \sigma_i(s) ds du + \sigma_i(t) \lambda_i t$$

The volatility functions $\sigma_i(t)$, t [0, T], i=1,…, N and market prices of risk λ_i, i=1,…, N satisfy the usual requirements and the N elements of the vector process $W=\{W_1(t),\ldots, W_N(t), t$ [0, T]$\}$ are independent Standard Brownian Motions, initialized at zero.

To apply a principal component analysis, the distribution at time t given the information at time 0 of an N-dimensional, vector of logarithms of bond prices is first derived, that is

$$\ln \boldsymbol{P}(t) = (\ln P(t, \tau_1, \ldots, \ln P(t, \tau_N))^T \qquad (9.2)$$

with $0 < \tau_1 < \ldots < \tau_N$. The reason for concentrating this analysis on the logarithms of bond prices instead of the familiar yields-to-maturity is mainly for expository purposes, as will become clear later on. This vector of logarithms of bond prices ln$\boldsymbol{P}(t)$ is normally distributed with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Omega}$:

$$\ln \boldsymbol{P}(t) \mid \mathcal{F}_0 \sim N(\boldsymbol{\mu}, \boldsymbol{\Omega})$$

The N-dimensional vector of means $\boldsymbol{\mu}$ is a function of the volatility functions, the vector of N market prices of risk, and the different times-to-maturity. The covariance matrix $\boldsymbol{\Omega}$ can simply be written as

$$\boldsymbol{\Omega} = tAA^T \qquad (9.4)$$

with

$$A = (a_{ij})_{i,j=1,\dots,N} = \left(\int_t^{t+\tau_i} \sigma_j(s)ds\right)_{i,j=1,\dots,N} \tag{9.5}$$

Alternatively, the vector ln$\boldsymbol{P(t)}$ can be written as a linear function of its N principal components[3] $\boldsymbol{x}$,

$$\ln \boldsymbol{P}(t) = \boldsymbol{\mu} + t^{\frac{1}{2}}\boldsymbol{\Gamma x} \tag{9.6}$$

with

$$\boldsymbol{\Gamma} = (\gamma_{ij})_{i,j=1\dots,N}$$

$$\boldsymbol{\Gamma}^T\boldsymbol{\Gamma} = \boldsymbol{I}$$

and

$$E(\boldsymbol{x} \mid \mathcal{F}_0) = 0$$

$$E(\boldsymbol{xx}^T \mid \mathcal{F}_0) = \boldsymbol{\Lambda}$$

$$\boldsymbol{\Lambda} = diag(\lambda_1, \dots, \lambda_N) \quad \lambda_i > \lambda_{i+1} \quad \lambda_n \geq 0$$

In terms of standardized components $\boldsymbol{y}$ having unit variances, the above relationship becomes

$$\ln \boldsymbol{P}(t) = \boldsymbol{\mu} + t^{\frac{1}{2}}\boldsymbol{Ay} \tag{9.7}$$

implying the following relationship for the matrix of volatility functions

$$\boldsymbol{A} = \boldsymbol{\Gamma\Lambda}^{\frac{1}{2}} \tag{9.8}$$

To investigate the impact of the standardized components $\boldsymbol{y}$ on option valuation, first a presentation will be made of the familiar pricing formula of a European call option $C(K, \tau_i, \tau_j)$ with exercise price K and maturity τ_j on a discount bond with τ_i at time 0,

$$C(K, \tau_i, \tau_j) = P(0, \tau_i)N(d_1) - KP(0, \tau_j)N(d_2) \tag{9.9}$$

with

$$d_1 = \frac{1}{v} \ln\left(\frac{P(0, \tau_i)}{KP(0, \tau_j)}\right) + \frac{1}{2} v$$

$$d_2 = d_1 - v$$

$$v^2 = \tau_j \sum_{k=1}^{k=N} \left(\int_{\tau_j}^{\tau_i} \sigma_k(s) \right)^2$$

The volatility parameter v^2 can be rewritten as follows

$$v^2 = \tau_j \sum_{k=1}^{k=N} \left(\int_{\tau_j}^{\tau_i} \sigma_k(s))^2 = \tau_j \sum_{k=1}^{k=N} (a_{ik} - a_{jk} \right)^2 \qquad (9.10)$$

Finally,

$$v^2 = \tau_j \sum_{k=1}^{k=N} (\gamma_{ik} - \gamma_{jk})^2 \lambda_k \qquad (9.11)$$

Principal component analysis, now, is selecting a certain number of factors A4 such that a given fraction α of the total variance of the elements of the vector ln***P(t)***, is explained, that is:

$$\frac{\sum_{i=1}^{i=M} \lambda_i}{\sum_{i=1}^{i=N} \lambda_i} \geq \alpha \qquad (9.12)$$

However, even if λ_i, $i=M+1,\ldots, N$ is relatively small, the difference between elements of the corresponding eigenvector can be relatively large, making the product of the two a significant part of the volatility parameter v^2 of the option. Because this can also happen the other way around, the first M factors do not necessarily explain a significant part of v^2. As mentioned in the introduction to this chapter, principal component analysis selects a number of factors sufficient to describe or represent a preselected fraction of the total variability of the yield curve. However, certain options are dependent upon the variability of a very specific maturity segment of the term structure of interest rates not covered by the reduced number of factors.

Although the effect of a reduction of the number of factors can be quantified precisely in the case of European call and put options on discount bonds, it is not the case when valuing contingent claims, such as American call and put options, for which closed-form solutions are not available. The use of principal component analysis, therefore, is generally not justified.

THE HEATH, JARROW AND MORTON II MODEL

In the N-factor discrete time version of the model of Heath *et al.* (1990b), discussed in Chapters 5 and 6, the continuous time process of the forward rate curve is approximated by a discrete time process $\hat{f} = \{\hat{f}(t_j, s)$, $tj=j\Delta t$, $j=0,\ldots, n$, $s \quad [t_{j+1}, T]\}$, which obeys the following stochastic difference equation

$$\hat{f}(t_{j+1}, s) = \hat{f}(t_j, s) + \sum_{i=1}^{i=N} \delta_i(t_j, s)\Delta t + \sum_{i=1}^{i=N} \sigma_i(t_j, s)(W_i(t_{j+1}) - W_i(t_j)) \qquad (9.13)$$

The vector process $\{W_1(t_{j+1}) - W_1(t_j), \ldots, W_N(t_{j+1}) - W_N(t_j), j=0,\ldots, n-1\}$ contains N independently binomial distributed random variables that can either increase or decrease at some time by $\sqrt{\Delta t}$ with equal probability. The function $\delta_i(t_j, s)$, $i=1,\ldots, N$, $j=0,\ldots, n-1$, is defined as

$$\delta_i(t_j, s) = \alpha_i(t_j, s) + \lambda_i t_j \qquad (9.14)$$

with λ, $i= 1,\ldots, N$, the market price of risk and $\alpha_i(t_j, s)$ defined according to equation (6.61). The corresponding volatility function $\sigma_i(t_j, s)$ is equal to

$$\sigma_i(t_j, s) = \sigma_i(s - t_j) Min(f_{Max}, \hat{f}(t_j, s)) \qquad (9.15)$$

with f_{Max} a real-valued constant and $\sigma_i(s - t_j)$ a deterministic function dependent on the remaining time-to-maturity $s - t_j$ of the forward rate.

Because closed-form solutions for the value of contingent claims cannot be derived in this "almost-proportional" interest rate model of Heath *et al.*, a numerical example using the techniques of Chapter 6 (pp. 132–5) is necessary to illustrate the impact of principal component analysis on derivative security valuation. For this purpose, a three-factor economy is assumed and the matrix A is defined as follows

$$\Omega = AA^T \qquad (9.16)$$

with

$$A = (a_{ij})_{i,j=1,2,3} = (\sigma_j(i+1))_{i,j=1,2,3} \qquad (9.17)$$

The particular elements of the matrix A are given by

$$\begin{pmatrix} 0.516 & 0.3 & 0.058 \\ 0.516 & 0 & -0.115 \\ 0.516 & -0.3 & 0.058 \end{pmatrix}$$

The eigenvalues or variances of the three factors are 0.80, 0.18 and 0.02 respectively, which means that the third factor counts for only 2 per cent of the total variability of the yield curve over time. The trading interval [0, T] is assumed to be divided into four periods of unit length, implying $T = 4$. The forward rate curve at the initial trade date 0, $f(0, 0)$, $f(0, 1)$, $f(0, 2)$ and $f(0, 3)$ is upward sloping with corresponding values 0.10, 0.12, 0.14 and 0.16, respectively.

In Table 9.1, the numerically obtained European call option values for different option and bond maturities are shown as a function of an increasing number of factors. The exercise price of the option is equal to the forward price of the underlying bond plus 10 per cent, while the face of the underlying discount bond is 100. From this table, similar evidence regarding the appropriateness of the application of principal component analysis to bond option valuation is obtained. Although a reduced number of factors represents 98 per cent of the total variability of the term structure of interest rates over time, the particular dependence of option values on this variability or volatility is not well described. Based on three factors, for example, the value of a one-year call European call option of a two-year bond is 7.7 cents. However, using the first two factors, the value of this option is only 0.2 cents.

Table 9.1 European call option values

	Option maturity	*Bond maturity*		
		2	*3*	*4*
One factor		0	1.167	3.375
Two factors	1	0.002	1.295	3.375
Three factors		0.077	1.319	3.375
One factor			0.109	2.020
Two factors	2		0.277	2.326
Three factors			0.332	2.373
One factor				0
Two factors	3			0
Three factors				0.004

This table contains the numerically obtained European call option values in the case of the model of Heath, Jarrow and Morton (1990b) II as a function of the different factors for different option and bond maturities. The exercise price of an option is equal to the forward price of the underlying bond plus 10 per cent.

Although the numerical example in this section is explicitly designed to illustrate the basic problem of the application of a principal component analysis to represent the volatility of the term structure of interest rates over time, a general application of this technique to value complicated interest rate contingent claims is not allowed. Due to numerical limitations, the dependence of some securities on particular factor volatilities cannot be quantified and some serious undervaluation of the “true” value of the security may be the result.

NOTES

1 This chapter is based on “Note on the Interest Rate Contingent Claim Valuation and the Use of Principal Components”, J.F.J. de Munnik, 1994, The Review of Futures Markets, Vol. 13, No. 2, pp. 695–702.

2 For more details on the *N*-factor case, see Chapter 5, n. 28.

3 See Lawley and Maxwell (1971, pp. 6–17).

10
CONCLUSIONS AND FURTHER RESEARCH

The theoretical overview of the different interest rates models and the empirical implementation of some particular models reveal the academic attention that has been paid to the problem of the valuation of interest rate derivative securities and show that a lot of empirical research has yet to be carried out to evaluate and assess the different models.

The theoretical overview by means of some general characteristics of the interest rates models enables and facilitates the explanation of the different features of the models presented. In addition, it clearly shows the trade-off between the need to incorporate the different institutional and empirically observed characteristics of the yield curve and the corresponding complexity of the resulting model. The numerical approaches discussed provide a general framework for the actual valuation of interest rate derivative securities given some interest rates model. The institutional characteristics of commonly traded contingent claims do not generally allow for the derivation of so-called closed-form solutions, implying the necessity of accurate and stable numerical methods to approximate the claims' value and assess the sensitivities of the claims with respect to the different input parameters.

Results presented in the empirical part point to the conclusion that an empirical comparison between theoretically correct but complex models and relatively simple implementable models is necessary. The empirically obtained results from the Dutch Government Bond Market during the years 1989 and 1990 do not favor a particular model in which interest rates are positive in comparison to another model in which interest rates are normally distributed. In addition, it has been shown that statistical criteria to obtain a sufficiently accurate description of the stochastic dynamics of the yield curve may lead to serious pricing errors in some derivative securities, stressing the need to integrate the estimation of interest rate models and the valuation of interest rate derivative securities.

CONCLUSIONS

This section presents more specific conclusions with respect to the research objectives formulated in Chapter 1.

The problem of the valuation of interest rate derivative securities relies heavily on conditions regarding the stochastic dynamics of the underlying values of the claim or the term structure of interest rates to be able to exclude arbitrage opportunities and to derive the value of any claim. The first chapter of the theoretical part presented an extensive overview regarding these conditions in discrete-time and continuous-time economies.

Given the stochastic processes of the assets, riskless arbitrage opportunities are excluded if there exists a unique equivalent probability measure such that the values of the different assets relative to a short-term money market account are martingales. In this case, the security market is complete, implying that every contingent claim has a unique and arbitrage-free value equal to the discounted expected value of the specific payout of the claim under this martingale measure. It is important to realize that in the actual economy individual investors may have subjective probability beliefs regarding the expected returns on different assets. As long as their probability measures are equivalent and as long as there exists a unique equivalent martingale measure, riskless arbitrage opportunities do not exist and every contingent claim can be uniquely priced.

The many different models designed to value interest rate derivative securities, which are discussed and explained in Chapters 4 and 5, illustrate the academic interest during the last decade. In order to determine the theoretical and empirical circumstances under which a particular model should be preferred and to be able to make some suggestions for further research, it is necessary to classify the different models according to some general characteristics. The most important distinction between the different approaches is the modelling of the underlying values of the contingent claims.

The direct approach basically extends the Black-Scholes security market model to incorporate the institutional characteristics of bonds. Given the stochastic processes of the underlying values explicitly, the existence of a unique equivalent probability measure excludes riskless arbitrage opportunities between these underlying values and between these underlying values and the contingent claim. The possibility of arbitrage opportunities between other interest rate dependent assets is not taken into account and as such, the direct approach can also be classified as a partial equilibrium approach. The model presented is a combination of two existing models and is applicable to the valuation of options on pure discount bonds. Both the explicit modelling of the stochastic processes of coupon-paying bonds and the necessity of obtaining conclusive results regarding the exclusion of arbitrage, seem impossible.

In the indirect approach, all interest rate derivative securities are considered to be functions of the instantaneous rate of interest. To obtain the actual values of these securities, arbitrage opportunities between all securities have to be excluded. The existence of a unique equivalent martingale measure enables the derivation of the term structure of interest rates at some valuation date and establishes a general equilibrium between all interest rate dependent securities. The different models within this indirect class can be further classified according to the endogenous or exogenous specification of the yield curve.

Within the class of endogenous term structure of interest rates models, the drift and volatility functions of the stochastic process of the instantaneous short-term rate of interest are not functions of calendar time. Because the arbitrage-free values of discount bonds are determined by the discounted expectation of the final payment under the martingale measure, the yield curve at some valuation date is a function of the parameters of the stochastic process of this instantaneous spot rate in a risk-neutral economy. Some well-known models in this class are the Vasicek (1977) model, in which interest rates are normally distributed and mean reverting, and the Cox, Ingersoll and Ross (1985) model, in which the instantaneous variance of the spot rate is proportional to the value of the spot rate and interest rates are mean reverting.

By allowing the drift term of the stochastic process of the instantaneous spot rate to be a function of calendar time, the exogenous term structure of interest rates models enables the implementation of an exogenously estimated or observed term structure of interest rates. A well-known model within this class is that of Ho and Lee (1986), in which interest rates are normally distributed.

In addition to this distinction, the term structure of interest rates models can be further and similarly classified according to the endogenous and exogenous specification of the term structure of interest rate volatilities. The Heath, Jarrow and Morton (1990b) II model is an example of an interest rate model in which both term structures are exogenously specified.

The different numerical methods necessary to value an interest rate contingent claim, given some specific model, generally approximate the continuous-time stochastic process of the interest rate by a discrete-time process or interest rate tree. Dependent upon the recombining of the different paths of this tree, the tree is denoted as path-dependent or pathindependent. In the case where the drift and volatility functions of the stochastic process are not dependent upon the particular path followed by the interest rate, the approaches of Nelson and Ramaswamy (1990) and Tian (1991) binomially approximate the continuous-time process. In the method of Hull and White (1990a), however, this approximation is established by means of a trinomial interest rate tree. By changing the corresponding probabilities at each node, one can ensure convergence of the discrete-time distribution towards its continuous-time counterpart.

To assess the computational efficiency and numerical accuracy of these three different methods, values of European call and put options on discount bonds have been calculated for the Cox, Ingersoll and Ross (1985) model for several combinations of the input parameters. For each of the scenarios, a benchmark value has been calculated that can be regarded as the true value of the option. Based on this benchmark value, the computational efficiency in relation to the numerical accuracy can be determined by calculating for each of the methods the maximum number of nodes for which the difference between the corresponding option value and benchmark value is more than one cent. For a number of scenarios considered, the method of Tian (1991) performs best in terms of both the mean as well as the standard deviation.

Although the original discrete-time derivation of the Heath, Jarrow and Morton (1990a) I model results in an interest rate tree that is path-dependent, a transformation can be applied to their interest rate process after which the construction of a simple path-independent interest rate tree is possible. It has been shown that the resulting numerical algorithm to value contingent claims is directly implementable and computationally efficient. In addition, the transformation allows for the direct calculation of the sensitivities of the value of the contingent claim with respect to term structure movements or changes in the underlying values, facilitating practical applications.

In the empirical part, the models of Vasicek (1977) and Cox, Ingersoll and Ross (CIR) (1985) have been estimated by means of a time series and cross-sectional analysis. Apart from the interesting comparison of two different models within the class of endogenous term structure of interest rate models, two ways of estimation are possible, that allow for a comparison between an implicitly and explicity estimated interest rate process.

Based on an exact discrete-time representation of the interest rate processes, the Vasicek model is nested within the CIR model. Only for the six-month and twelve-month

AIBOR during the years 1985 through 1990, the Vasicek model has been rejected in favor of the CIR model. Based on a sample of liquid Dutch Government Bonds during 1989 and 1990, it can be concluded that the differences between the weekly estimated term structures of interest rates and the resulting implications for contingent claim pricing are small. Although further research should be carried out to investigate this conclusion, the results seem to suggest that the relatively simple model of Vasicek is able to capture the stochastic dynamics of interest rates in relation to the valuation of interest rate derivative securities in comparison to the more complicated CIR model.

The two estimation methodologies, however, do reveal significant differences between the cross-sectional and time series estimation of the stochastic process of the instantaneous short term rate of interest. The instantaneous volatility of the spot rate is significantly higher based on a time series analysis for both models. Based on a cross-sectional analysis, however, interest rates are close to a random walk and almost deterministic.

A principal component analysis to determine a relatively small number of factors to describe the stochastic dynamics of the yield curve over time is generally not applicable to value interest rate derivative securities. A theoretical and numerical investigation in case of the Jamshidian (1990) and Heath, Jarrow and Morton (1990b) II model, respectively, shows that, although a small number of factors sufficiently describe the general movement of the term structure of interest rates, the stochastic dynamics of specific maturity segments upon which a contingent claim critically depends may be ignored, thereby seriously underestimating the true value of the claim. Because the exact relationship between the stochastic dynamics of specific segments and the value of a contingent claim is generally unknown, the application of principal component analysis to value interest rate derivative securities is disputable.

FURTHER RESEARCH

As the last decade has shown a lot of theoretical progress with respect to the theoretical valuation of interest rate derivative securities, it can be expected that much attention will be paid to an empirical investigation of the different models.

Preliminary results regarding two specific models within the class of endogenous term structure of interest rate models reveal that the implied and explicit estimation of the stochastic process of the interest rate is significantly different. Although one may conclude that the models are misspecified, it is interesting to investigate an integration of the two different estimation methodologies and of the corresponding two different samples of observations.

In addition to an extension of the empirical research regarding endogenous term structure of interest rate models, future research should be focused on an empirical examination of the exogenous term structure of interest rate models. The explicit modelling of the stochastic evolution of the yield curve over time, given an observed yield curve at some valuation date, might be able to capture or describe the observed interest rate process more accurately than do the models within the class of endogenous term structure of interest rate models.

The last chapter of the empirical part has shown that applying principal component analysis to the stochastic movement of the yield curve is generally not justified to value interest rate derivative securities. First, it would be interesting to investigate the determination of an extended statistical criterium that not only describes the proportion explained by each of the different factors but also enables the assessment of the possible pricing errors of representative contingent claims. Second, it is necessary to examine the possibility of the existence of similar pricing errors resulting from estimating well-known models like the Vasicek model, the CIR model and the Heath, Jarrow and Morton I model. The cross-sectional estimation of the Vasicek model, for example, results implicitly in the stochastic description of different maturity segments. The specific stochastic characteristics of a particular segment may not be sufficiently represented as a result of the used statistical criteria. Further research should also be concentrated on a theoretical assessment and empirical investigation of this phenomenon.

REFERENCES

Amin, K., 1991, "On the Computation of Continuous Time Option Prices Using Discrete Approximations," *Journal of Financial and Quantitative Analysis,* Vol. 26, pp. 477–97.

Arnold, L., 1974, "Stochastic Differential Equations," John Wiley & Sons, Inc., New York.

Ball, C.A. and W.N. Torous, 1983, "Bond Price Dynamics and Options," *Journal of Financial and Quantitative Analysis,* Vol. 18, pp. 517–31.

Black, F., E. Derman, and W. Toy, 1990, "A One-Factor Model of the Term Structure of Interest Rates and Its Application to Treasury Bond Options," *Financial Analysts Journal,* Jan—Feb, pp. 33–9.

Black, F. and M. Scholes, 1973, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy,* Vol. 81, pp. 637–59.

Boyle, P.P., 1977, "Options: a Monte Carlo Approach," *Journal of Financial Economics,* Vol. 4, pp. 323–38.

Boyle, P.P., 1988, "A Lattice Framework for Option Pricing with Two State Variables," *Journal of Financial and Quantitative Analysis,* Vol. 23, pp. 1–12.

Boyle, P.P., J. Evnine and S. Gibbs, 1989, "Numerical Evaluation of Multivariate Contingent Claims," *Review of Financial Studies,* Vol. 2, pp. 241–51.

Brennan, M.J., 1979, "The Pricing of Contingent Claims in Discrete Time Models," *The Journal of Finance,* Vol. 24, No. 1, pp. 53–68.

Brennan, M.J. and E.S. Schwartz, 1983, "Alternative Methods for Valuing Debt Options," *Finance,* Vol. 4, pp. 119–37.

Brown, R.H. and S.M. Schaefer, 1994, "The Term Structure of Real Interest Rates and the Cox, Ingersoll and Ross Model," *Journal of Financial Economics,* Vol. 35, pp. 3–42.

Brown, S.J. and Ph.H. Dybvig, 1986, "The Empirical Implications of the Cox, Ingersoll and Ross Theory of the Term Structure of Interest Rates," *The Journal of Finance,* Vol. 41, pp. 617–30.

Bühler, W. and J. Käsler, 1989, "Konsistente Anleihepreise und Optionen auf Anleihen," *Unpublished Manuscript, Universität Dortmund.*

Chan, K.C., G.A. Karolyi, F.A. Longstaff and A.B. Sanders, 1991, "The Volatility of Japanese Interest Rates: A Comparison of Alternative Term Structure Models," *Unpublished Manuscript,* Ohio State University.

Chan, K.C., G.A. Karolyi, F.A. Longstaff and A.B. Sanders, 1992, "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate", *The Journal of Finance* , Vol. 47, pp. 1209–27.

Cheng, S.T., 1989, "On the Feasibility of Arbitrage-Based Option Pricing When Stochastic Bond Price Processes Are Involved," *Journal of Economic Theory,* Vol. 53, pp. 185–98.

Chvátal, V., 1983, "Linear Programming," W.H. Freeman and Co., New York.

Courtadon, G., 1982, "The Pricing of Options on Default-Free Bonds," *Journal of Financial and Quantitative Analysis,* Vol. 17, pp. 75–101.

Cox, J.C., J.E. Ingersoll and S.A. Ross, 1985, "A Theory of Term Structure of Interest Rates," *Econometrica,* Vol. 53, pp. 385–408.

Dothan, L.U., 1978, “On the Term Structure of Interest Rates,” *Journal of Financial Economics,* Vol. 6, pp. 59–69.
Duan, J.-C., 1990, “The Garch Option Pricing Model,” *Unpublished Manuscript,* McGill University, Montreal.
Duffie, D., 1988, “Security Markets, Stochastic Models,” Academic Press, Inc., San Diego, California.
Dybvig, Ph.H., 1980, “A Positive Wealth Constraint Precludes Arbitrage Profits (e.g. from Doubling) in the Black-Scholes model,” *Unpublished Manuscript,* Princeton University, New Jersey.
Dybvig, Ph.H., 1989, “Bond and Bond Option Pricing Based on the Current Term Structure,” *Unpublished Manuscript,* Washington University in Saint Louis.
Elliott, R.J., 1982, “Stochastic Calculus and Applications,” Springer-Verlag, New York.
Feller, W., 1951, “Two Singular Diffusion Problems,” *Annals of Mathematics,* Vol. 54, pp. 173–82.
Fuller, W.A., 1976, “Introduction to Statistical Time Series,” Wiley, New York.
Garbade, K.D., 1986, “Modes of Fluctuation in Bond Yields—An Analysis of Principal Components,” *Unpublished Manuscript,* Money Market Center, Bankers Trust Company.
Geske, R. and K. Shastri, 1985, “Valuation by Approximation: A Comparison of Alternative Valuation Techniques,” *Journal of Financial and Quantitative Analysis,* Vol. 20, No.l, pp. 45–71.
Gihman, I.I. and A.V. Skorohod, 1972, “Stochastic Differential Equations,” Springer-Verlag, New York.
Harrison, J.M. and D. Kreps, 1979, “Martingales and Arbitrages in Multiperiod Security Markets,” *Journal of Economic Theory,* Vol. 20, pp. 381–408.
Harrison, J.M. and S.R. Pliska, 1981, “Martingales and Stochastic Integrals in the Theory of Continuous Trading,” *Stochastic Processes and their Applications,* Vol. 11, pp. 215–360.
Harvey, A.C., 1990, “The Econometric Analysis of Time Series,” Philip Allan.
Heath, D.C. and R.A. Jarrow, 1987, “Arbitrage, Continuous Trading, and Margin Requirements,” *The Journal of Finance,* Vol. 42, No. 5, pp. 1129–42.
Heath, D.C., R.A. Jarrow and A. Morton, 1990a, “Bond Pricing and the Term Structure of Interest Rates: A Discrete Time Approximation,” *Journal of Financial and Quantitative Analysis,* pp. 419–40.
Heath, D.C., R.A. Jarrow and A. Morton, 1990b, “Contingent Claim Valuation with a Random Evolution of Interest Rates,” *The Review of Futures Markets,* Vol. 9, pp. 54–76.
Heath, D.C., R.A. Jarrow and A. Morton, 1992, “Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation,” *Econometrica,* Vol. 60, pp. 77–105.
Ho, T.S.Y, and S.-B. Lee, 1986, “Term Structure Movements and Pricing Interest Rate Contingent Claims,” *The Journal of Finance,* Vo1. 41, pp. 1011–29.
Hull, J., 1989, “Options, Futures, and Other Derivative Securities,” Prentice-Hall International Editions.
Hull, J. and A. White, 1988, “The Use of Control Variate Technique in Option Pricing,” *Journal of Financial and Quantitative Analysis,* Vol. 23, pp. 237–51.
Hull, J. and A. White, 1990a, “Valuing Derivative Securities Using the Explicit Finite Difference Method,” *Journal of Financial and Quantitative Analysis,* Vol. 25, pp. 87–100.
Hull, J. and A. White, 1990b, “Pricing Interest-Rate-Derivative Securities,” *The Review of Financial Studies,* Vol. 3, pp. 573–92.
Ingersoll, Jr., J.E., 1987, “Theory of Financial Decision Making,” Rowman & Littlefield studies in financial economics.
Jamshidian, F., 1989, “The Multifactor Gaussian Interest Rate Model and Implementation,” *Unpublished Manuscript,* Merrill Lynch Capital Markets, New York.
Jamshidian, F., 1990, “Bond and Option Valuation in the Gaussian Interest Rate Model,” *To appear in Research in Finance,* Vol. 9.

Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lütkepohl and T.-C. Lee, 1982, “Introduction to the Theory and Practice of Econometrics,” John Wiley & Sons, Inc., New York.

Kendall, M. and A. Stuart, 1979, “Advanced Theory of Statistics,” Vol. 2, Macmillan, New York.

Kreps, D., 1979, “Three Essays on Capital Markets,” *Technical Report No. 298,* Institute for Mathematical Studies in the Social Sciences.

Langetieg, T.C., 1980, “A Multivariate Model of the Term Structure,” *The Journal of Finance,* Vol. 35, pp. 71–97.

Lawley, D.N. and A.E. Maxwell, 1971, “Factor Analysis as a Statistical Method,” Butterworths, London.

Litterman, R. and J. Scheinkman, 1988, “Common Factors affecting Bond Returns,” *Unpublished Manuscript,* Goldman Sachs.

Lo, A.W., 1986, “Statistical Tests of Contingent-Claims Asset-Pricing Models,” *Journal of Financial Economics,* Vol. 17, pp. 143–73.

Maddala, G.S., 1988, “Introduction to Econometrics,” Macmillan, New York.

Malkiel, B.G., 1990, “A Random Walk Down Wall Street,” W.W. Norton & Company, New York.

Merton, R.C., 1973, “Theory of Rational Option Pricing,” *Bell Journal of Econom-ics and Management Science,* Vo1. 4, pp. 141–83.

Morton, A., 1988, “A Class of Stochastic Differential Equations Arising in Models for the Evolution of Bond Prices,” *Unpublished Manuscript,* School of Operations Research and Industrial Engineering, Cornell University.

Müller, S., 1985, “Arbitrage Pricing of Contingent Claims,” Lecture Notes in Economics and Mathematical Systems, Vol. 254, Springer-Verlag, New York.

de Munnik, J.F.J., 1992, “The Construction of a Path-Independent Interest Rate Tree: The Model of Heath, Jarrow and Morton,” *Advances in Futures and Options Research,* Vol. 7, pp. 135–45.

de Munnik, J.F.J., 1994b, “Note on the Interest Rate Contingent Claim Valuation and the Use of Principal Components,” *The Review of Futures Markets,* Vol. 13, No. 2, pp. 695–702.

de Munnik, J.F.J. and A.C.F. Vorst, 1988, “De Waardering van Vervroegd Aflosbare Staatsobligaties,” (Trans.:“The Valuation of Callable Bonds”) in: “Finan-ciering en Belegging, Stand van Zaken anno 1988", H. Berkman, J. van der Meulen and A.B.M. Soppe, pp. 355–66.

de Munnik, J.F.J. and P.C. Schotman, 1994, “Cross Sectional versus Time Series Estimation of Term Structure Models: Empirical Results for the Dutch Bond Market,” *Journal of Banking and Finance,* Vol. 18, No. 5, pp. 997–1025.

Musiela, M., S.M. Turnbull and L.M. Wakeman, 1990, “Interest Rate Risk Management,” *Unpublished Manuscript,* School of Business, Queen’s University.

Nelson, D.B. and K. Ramaswamy, 1990, “Simple Binomial Processes as Diffusion Approximations in Financial Models,” *Review of Financial Studies,* Vol. 3, pp. 393–430.

Pedersen, H.W., E.S.W. Shiu and A.E. Thorlacius, 1989, “Arbitrage-Free Pricing of Interest-Rate Contingent Claims,” *Transactions of the Society of Actuaries,* Vol. 41, pp. 231–66.

Rady, Sv. and K. Sandmann, 1992, “The Direct Approach to Debt Option Pricing,” *Unpublished Manuscript,* Rheinische Friedrich-Wilhelms Universität Bonn.

Redington, F.M., 1952, “Review of the Principles of Life-Office Valuations,” *Journal of the Institute of Actuaries,* Vol. 18.

Richard, S.F., 1978, “An Arbitrage Model of the Term Structure of Interest Rates,” *Journal of Financial Economics,* Vo1. 6, p. 33–57.

Schaefer, S.M. and E.S. Schwartz, 1987, “Time-Dependent Variance and the Pricing of Bond Options,” *The Journal of Finance,* Vol. 42, No. 5, pp. 1113–28.

Schöbel, R., 1986, “On a Difference Between Options on Stocks and Default-free Debt Options,” Arbeitsbericht des Fachbereichs Wirtschafts- und Sozial-wissenschaften.

Tian, Y., 1991, “A Simplified Binomial Approach to the Pricing of Interest-Rate Contingent Claims,” *Unpublished Manuscript,* Faculty of Administrative Studies, York University.

Vasicek, O., 1977, “An Equilibrium Characterization of the Term Structure,” *Journal of Financial Economics,* Vol. 5, pp. 177–88.

INDEX