Efficient Methods For Valuing Interest Rate Derivatives

Book · January 2000

DOI: 10.1007/978-1-4471-3888-4

Antoon Pelsser
Maastricht University

About the Author

Antoon Pelsser is currently working with the ABN-Amro bank, in the department Derivatives Special Trading and Research at the head-office in Amsterdam. His current research interests include building and developing models for pricing and managing interest rate derivatives. Previously he worked at the Tinbergen Institute in Rotterdam, where he investigated the influence of transaction costs on the prices of financial instruments. He has published in several journals including European Journal of Operational Research and the Journal of Derivatives. Mr. Pelsser holds a degree in Econometrics from Erasmus University Rotterdam.

Acknowledgements

During the period I was working on this thesis, I received help and guidance from a lot of different people.

First, I would like to thank my mentors Ton Vorst and Douglas Bongartz-Renaud. Ton has shown me how to be a scientist, Doug has shown me how to be a practitioner. Having contact to both a scientist and a practitioner has proven to be very fruitful. Their ideas and thoughts can be found throughout this thesis, and without their help and inspiration I would never have completed this project.

Also, I would like to thank my colleagues Carien Dam and Boris van der Saag. They have patiently endured my desire to share with them every new idea I had. They also read and commented an endless stream of papers and chapters I dropped on their desks.

Furthermore, I would like to thank Frank de Jong. Unlike me, he is a true econometrician and I was very glad I could make use of his expertise. He has put me on the right track in the area of non-linear and non-nested hypothesis testing.

Last, but not least, I would like to express my gratitude towards my family and friends who have encouraged and supported me while I embarked on this strange project which had "something to do with options". My mother has, from my childhood onwards, always convinced me that I could achieve anything I wanted, and this thesis shows that she was right. Special thanks go to Chantal who has continually supported me during this past period.

The research in this thesis was partially supported by the Netherlands Organisation for Scientific Research (NWO) and by the Erasmus Center for Financial Research (ECFR).

Contents

1

Introduction

SINCE the opening of the first options exchange in Chicago in 1973, the financial world has witnessed an explosive growth in the trading of derivative securities. Since that time, exchanges where futures and options can be traded have been opened all over the world and the volume of contracts traded worldwide has grown enormously.

The growth in derivatives markets has not only been a growth in volume, but also a growth in complexity. Most of these more complex derivative contracts are not exchange traded, but are traded "over-the-counter". Often, the over-the-counter contracts are created by banks to provide tailor made products to reduce financial risks for clients. In this respect, banks played an innovative and important role in providing a market for the exchange of financial risks.

However, derivatives can also be used to create highly leveraged speculative positions, which can lead to large profits, or large losses. In recent years, several companies have suffered large losses due to speculative trading in derivatives. In some cases these losses made the headlines of the financial press. This has created a general feeling that "derivatives are dangerous". Some people have called for a strict regulation of derivatives markets, or even for a complete ban of over-the-counter trading.

Although these claims are an over-reaction to what has recently happened, it has become apparent that it is important, both for market participants and regulators, to have a good insight into the pricing and risk characteristics of derivatives. Considerable academic research has been devoted to the valuation of derivative securities. Since the seminal paper of Fisher Black and Myron Scholes in 1973, an elegant theory has been developed.

The cover of this thesis shows three interconnected rings which are labelled Arbitrage, Martingales and Partial Differential Equations. An *arbitrage opportunity* offers the chance of generating wealth at some future time without an initial investment. Due to competitive forces, it seems obvious that arbitrage opportunities cannot exist in an economy that is in equilibrium. This is often summarised by the phrase "there is no such thing as a free lunch." A *martingale* is the mathematical formalisation of the concept of a fair game. If prices of derivative securities can be modeled as martingales it means that no market participant can consistently make (or lose) money by trading in derivative securities.[1] *Partial differential equations* can be used to describe the dynamic behaviour of the prices of derivatives. We can borrow from physics many of the methods that have been developed to solve partial differential equations. It will provide us with important tools to explicitly calculate prices of derivatives.

The concepts of arbitrage, martingales and partial differential equations are the three pillars on which the theory of the valuation of derivative securities rests. A closer inspection of the rings will reveal that every pair of rings is not connected, only the three rings together are connected. The same is true for the theory of valuing derivative securities. A good command of the concepts of arbitrage, martingales and partial differential equations is needed to obtain a coherent understanding of the valuation of derivative securities. If either of the three concepts is neglected, the theory falls apart. Chapter 2, which explains the basic theory of the valuation of derivatives is therefore called "Arbitrage, Martingales and Partial Differential Equations".

After Chapter 2, the thesis is split in two parts. The first part is concerned with exploring efficient methods for valuing and managing interest rate derivatives. The second part of this thesis is concerned with efficient methods for other derivative securities such as stock options or foreign-exchange options.

[1] By introducing imperfections into the market (like transaction costs) some market participants can consistently make money. However, in the absence of any market imperfections trading in derivatives would really be a "fair game".

As was argued before, the markets in derivative securities can be viewed as insurance markets for financial risks. Since the Fed decided in 1979 to change its monetary policy, interest rate volatilities in the US have risen considerably. Due to the increasing globalisation of capital markets, this has led to an increase in interest rate volatilities world-wide. Many companies have sought to buy insurance against the increased uncertainty in interest rate markets. For this reason, the market for interest rate derivatives has been one of the fastest growing markets in the last decade. Strong interest in this area has inspired a lot of research into modeling the behaviour of interest rates and the pricing and risk characteristics of interest rate derivatives.

Two major types of modeling approaches can be distinguished. The first approach is to describe the spot interest rate by a stochastic differential equation with constant parameters. This leads to models which have an endogenous term-structure of interest rates. Due to the fact that the term-structure of interest rates is endogenously given in these models, it is not possible to fit the observed term-structure of interest rates (and hence the observed bond- or swap-prices) exactly. However, for models that are used to price and manage interest rate derivatives, it is imperative that the underlying securities (like bonds or swaps) are priced accurately, hence models with an endogenous term-structure have little practical interest or application.

This shortcoming has inspired the second type of approach, which takes the initial term-structure of interest rates as given, and describes the dynamics of the complete term-structure. In this thesis we will be concerned with this category of models only.

The first part of this thesis is organised as follows. In Chapter 3 we show how interest rate derivatives can be valued theoretically. Interest rates play a double role when valuing interest rate derivatives as they determine both the discounting and the payoff of the derivative. It turns out that interest rate derivatives are more difficult to value than other derivatives. Hence, we explore in Chapter 4 some analytical methods that can be used to simplify the pricing of interest rate derivatives. In Chapters 5 and 6 we analyse two interest rate models. Due to the assumption that there is only one underlying source of uncertainty that drives the evolution of the interest rates, these models are relatively simple to understand and to analyse. We derive analytical formulæ for valuing interest rate derivatives, and we derive numerical methods to approximate the prices of derivatives for which no analytical pricing formulæ can be found. Finally, we make an empirical comparison of several one-factor models in Chapter 7. Unfortunately, it turns out that the models with a rich analytical structure do not describe the prices of interest rate derivatives very well. However, an alternative model is proposed that fits the data better. This alternative model is analytically not very tractable and only numerical methods can be used to approximate prices of interest rate derivatives.

The second part of this thesis is concerned with efficient methods for valuing and hedging derivative securities like stock or foreign-exchange options. This second part is organised as follows.

First, we consider efficient methods to calculate hedging parameters from binomial trees. Binomial trees are widely used to calculate values for options (like American-style options) for which no analytic formulæ are available. Hence, binomial trees are a very important tool for practitioners. Many people who use binomial trees believe that calculating hedging parameters from nodes within the tree is a "quick and dirty" way of obtaining the hedging parameters delta, gamma and theta. It is generally thought that a finite difference method provides a slow but accurate answer. However, in Chapter 8, we demonstrate that the contrary is true: the calculation within the tree provides the most accurate answer.

In Chapter 9 we show how a portfolio manager can use options to reshape the return distribution of a portfolio at minimal initial costs. If one approximates the return distribution in a certain way, the problem at hand is a linear programming problem, which is relatively simple to solve. The methodology used is very general and can be extended in several ways. For example, the methodology can even be used to find an optimal portfolio in the presence of transaction costs.

In Chapter 10 we relax the assumption of the absence of transaction costs, which is made throughout this thesis. We investigate the efficiency of dynamic portfolio strategies in the presence of transaction costs. We show that dynamic portfolio strategies that are efficient when no transaction costs are present, but that require frequent rebalancing of the portfolio, are in the presence of transaction costs quickly dominated by very simple (and "inefficient") strategies like stop-loss or lock-in. We also show how, in the presence of transaction costs, options can be used to enhance the opportunity set of agents in an economy.

Finally, we summarise and conclude in Chapter 11.

2

Arbitrage, Martingales and Partial Differential Equations

THE cornerstone of option pricing theory is the assumption that any financial instrument which has a non-negative payout must have a non-negative price. The existence of an instrument which would have non-negative payoffs and a negative price is called an *arbitrage opportunity.* If arbitrage opportunities would exist, it would be a means for investors to generate money without any initial investment. Of course, many investors would try to exploit the arbitrage opportunity, and due to the increased demand, the price would rise and the arbitrage opportunity would disappear. Hence, in an economy that is in equilibrium it seems reasonable to rule out the existence of arbitrage opportunities. Although the assumption that arbitrage opportunities do not exist seems a rather plausible and trivial assumption, we shall see it is indeed the foundation for all of the option pricing theory.

Another important assumption needed to get the edifice of option pricing off the ground is the absence of *transaction costs.* This means that assets can be bought and sold in the market for the same price. This assumption is clearly violated in real markets. In the presence of transaction costs, not all arbitrage opportunities which would theoretically be profitable can be exploited. However, large market participants (like banks and institutions) face very little transaction costs. These large players have the opportunity to exploit almost all arbitrage opportunities with large amounts of money and markets will be driven to an equilibrium close to the equilibrium that would prevail if transaction costs were absent. Hence, if we consider markets as a whole, the assumption that transaction costs are absent is a good approximation of the real world situation.

Individual investors often face larger transaction costs when trading in financial markets, and they can often use options and other derivative securities to enlarge their opportunity set. We expand on this point in Part II of this thesis.

The rest of this chapter is divided in two sections. In the first section we provide the basic mathematical setup that will be used throughout this thesis. We furthermore show under which conditions an economy is free of arbitrage opportunities and how prices of derivative securities can be calculated. As an example we analyse the Black-Scholes model. In the second section we demonstrate how prices for derivatives can be calculated by solving partial differential equations. We show how a partial differential equation can be derived from no-arbitrage arguments, and we analyse analytical and numerical methods for solving partial differential equations.

1. Arbitrage and Martingales

In this section we provide the basic mathematical setting in which the theory of option pricing can be cast. We derive the key result that an economy is free of arbitrage opportunities if a probability measure can be found such that the prices of marketed assets become martingales. By setting up trading strategies which replicate the payoff of derivative securities, the martingale property can then be shown to carry over from the marketed assets to the prices of all derivative securities. Hence, the prices of all derivatives become martingales and this property can then be used to calculate prices for derivative securities.

1.1. Basic Setup

Throughout this thesis we consider a continuous trading economy, with trading interval given by $[0, T]$. The uncertainty is modeled by the probability space $(\Omega, \mathcal{F}, Q)$. In this notation, Ω denotes a sample space, with elements $\omega \in \Omega$; $\mathcal{F}$ denotes a σ-algebra on Ω; and Q denotes a probability measure on $(\Omega, \mathcal{F})$.[1] The uncertainty is resolved over $[0, T]$ according

[1] For a formal introduction to the concepts used in this section, see Rogers and Williams (1994); for a more intuitive introduction, see Duffie (1988) or Dothan (1990).

to a filtration $\{\mathcal{F}_t\}$ satisfying "the usual conditions", generated by a Brownian Motion[2] W initiated at 0.

Throughout this thesis we assume that there exist assets which are traded in a market. The assets are called *marketed assets.* We also assume that the prices $Z(t)$ of these marketed assets can be modeled via Itô processes which are described by stochastic differential equations

$$dZ(t) = \mu(t,\omega)\,dt + \sigma(t,\omega)\,dW, \tag{1}$$

where the functions $\mu(t,\omega)$ and $\sigma(t,\omega)$ are assumed to be $\mathcal{F}_t$-adapted and also satisfy

$$\begin{aligned} &\int_0^T |\mu(t,\omega)|\,dt < \infty \\ &\int_0^T \sigma(t,\omega)^2\,dt < \infty, \end{aligned} \tag{2}$$

with probability one.

The observant reader may note that there is only one source of uncertainty (the Brownian Motion W) that drives the prices of the marketed assets. It is relatively straightforward to set up the economy such that more sources of uncertainty drive the prices of the marketed assets. However, for the largest part of this thesis we will be concerned with economies which are assumed to have only one source of uncertainty.

It is also true that the prices of marketed assets defined in (1) are less general than usual in the literature. The sample paths of Itô processes are continuous, which excludes discrete dividend payments. It is possible to develop a general theory on the basis of right-continuous with left-limits (RCLL) processes[3] which allow for a countable number of discontinuities in the sample paths. However, in this thesis we will nowhere encounter marketed assets with discontinuous sample paths.

Suppose there are N marketed assets with prices $Z_1(t),\ldots,Z_N(t)$, which all follow Itô processes. A *trading strategy* is a predictable[4] N-dimensional stochastic process $\delta(t,\omega) = \big(\delta_1(t,\omega),\ldots,\delta_N(t,\omega)\big)$, where

---

[2] Brownian Motion is also called a *Wiener process.*

[3] RCLL processes are also known as *càdlàg*, *corlol* or *R-processes* in the literature.

[4] Predictable processes are also known as *previsible processes.*

$\delta_n(t, \omega)$ denotes the holdings in asset n at time t. The asset holdings $\delta_n(t, \omega)$ are furthermore assumed to satisfy an additional regularity condition to which we will return later.

The value $V(\delta, t)$ at time t of a trading strategy δ is given by

$$V(\delta, t) = \sum_{n=1}^{N} \delta_n(t) Z_n(t). \tag{3}$$

A *self-financing trading strategy* is a strategy δ with the property

$$V(\delta, t) = V(\delta, 0) + \sum_{n=1}^{N} \int_0^t \delta_n(s)\, dZ_n(s), \quad \forall t \in [0, T]. \tag{4}$$

Hence, a self-financing trading strategy is a trading strategy that requires nor generates funds between time 0 and time T.

An *arbitrage opportunity* is a self-financing trading strategy δ, with $\Pr[V(\delta, T) \geq 0] = 1$ and $V(\delta, 0) < 0$. Hence, an arbitrage opportunity is a self-financing trading strategy which has strictly negative initial costs, and with probability one has a non-negative value at time T.

A *derivative security* is defined as a $\mathcal{F}_T$-measurable random variable $H(T)$. The random variable has to satisfy an additional regularity constraint to which we will return later. The random variable $H(T)$ can be interpreted as the (uncertain) payoff of the derivative security at time T. If we can find a self-financing trading strategy δ such that $V(\delta, T) = H(T)$ with probability one, the derivative is said to be *attainable*. The self-financing trading strategy is then called a *replicating strategy*. If in an economy all derivative securities are attainable, the economy is called *complete*.

If no arbitrage opportunities and no transaction costs exist in an economy, the value of a replicating strategy at time t gives a unique value for the attainable derivative $H(T)$. This is true, since (in the absence of transaction costs) the existence of two replicating strategies of the same derivative with different values would immediately create an arbitrage opportunity. Hence, we can determine the value of derivative securities by the value of the replicating portfolios. This is called *pricing by arbitrage*.

However, this raises two questions. First, under which conditions is a continuous trading economy free of arbitrage opportunities? Second, under which conditions is the economy complete? If these two conditions are satisfied, all derivative securities can be priced by arbitrage.

1.2. Equivalent Martingale Measure

The questions of no-arbitrage and completeness were mathematically rigorously analysed in the seminal papers of Harrison and Kreps (1979) and Harrison and Pliska (1981). They showed that both questions can be solved at once, if we express prices in the economy in terms of a numeraire.

Any marketed asset which has strictly positive prices (and pays no dividends) for all $t \in [0, T]$ is called a *numeraire*. We can use numeraires to denominate all prices in an economy. Suppose that the marketed asset Z_1 is a numeraire. The prices of other marketed assets denominated in Z_1 are called the *relative prices* denoted by $Z'_n = Z_n / Z_1$.

Let $(\Omega, \mathcal{F}, Q)$ denote the probability space from the previous subsection. Consider now the set $\mathcal{Q}$ which contains all probability measures Q^* such that:

i Q^* is equivalent to Q, i.e. both measures have the same null-sets;
ii the relative price processes Z'_n are martingales under Q^* for all n, i.e. for $t \leq s$ we have $E^*(Z'_n(s) \mid \mathcal{F}_t) = Z'_n(t)$.

The measures $Q^* \in \mathcal{Q}$ are called *equivalent martingale measures*. Suppose we take one equivalent martingale measure $Q^* \in \mathcal{Q}$. Then, in terms of this "reference measure", we can give precise definitions for derivative securities and trading strategies given in the previous subsection.

A *derivative security* is a $\mathcal{F}_T$-measurable random variable $H(T)$ such that $E^{Q^*}\big(|H(T)|\big) < \infty$.[5] Hence, derivative securities are those securities for which the expectation of the payoff is well-defined.

A *trading strategy* is a predictable N-dimensional stochastic process $\big(\delta_1(t,\omega), \ldots, \delta_N(t,\omega)\big)$ such that the stochastic integrals

$$\int_0^t \delta_n(s)\, dZ'_n(s) \tag{5}$$

are martingales under Q^*. For self-financing strategies this implies that the value $V'(\delta, t)$ in terms of the relative prices Z' is a Q^*-martingale.

The condition on trading strategies is a rather technical condition. It arises from the fact that for predictable processes in general, the value

[5] The set of all random variables that satisfy these constraints is often denoted $\mathcal{L}^1(\Omega, \mathcal{F}_T, Q^*)$.

processes $V'(\delta, t)$ of self-financing trading strategies are only *local* martingales under Q^*. For a local martingale

$$\sup_{t \in [0,T]} \left\{ E^{Q^*} \left(V'(\delta, t) \right) \right\} = \infty \tag{6}$$

is possible, while for martingales

$$\sup_{t \in [0,T]} \left\{ E^{Q^*} \left(V'(\delta, t) \right) \right\} < \infty \tag{7}$$

is always satisfied. This difference between local martingales and martingales allows for the existence of so-called *doubling strategies*, which are arbitrage opportunities. This was first pointed out by Harrison and Pliska (1981). Hence, an economy can only be arbitrage-free if the value processes of self-financing trading strategies are martingales.

Several restrictions can be imposed on the processes δ to ensure the martingale property of the value processes $V'(\delta, t)$. For a discussion, see Harrison and Pliska (1981) or Duffie (1988). Heath and Jarrow (1987) and Dothan (1990) show that the presence of wealth constraints or constraints like margin requirements also ensures that the value processes are martingales. Because these constraints are actually present in security markets, it will be assumed throughout this thesis that this restriction holds.

Subject to the definitions given above, we have the following result:

Unique Equivalent Martingale Measure. *A continuous economy is free of arbitrage opportunities and every derivative security is attainable if $\mathcal{Q}$ contains only one equivalent martingale measure.*

For numeraires which are non-stochastic this result was proved by Harrison and Pliska (1981). For a general proof, see Dothan (1990), Chapter 12. ■

The definition of $\mathcal{Q}$ depends on the choice of numeraire. Hence, for a different choice of numeraire we obtain a different set $\mathcal{Q}$ of equivalent martingale measures and a different unique equivalent martingale measure. Hence, "unique" only means that the set $\mathcal{Q}$ contains only one element. To make things worse, the definition of derivative securities changes also with a different choice of numeraire. It is therefore conceivable that a payoff pattern $H(T)$ which can be replicated for one choice of numeraire, cannot be replicated for another choice of numeraire. Much works remains to be done in this area, for example in establishing which set of payoff patterns can be replicated for all choices of numeraire. However, for the derivatives we analyse in this thesis, we never encounter such a situation, and we will implicitly make the assumption that the payoff patterns we analyse can be replicated with any choice of numeraire.

From the result given above follows immediately that for a given numeraire Z with unique equivalent martingale measure Q_Z, the value of a self-financing trading strategy $V'(\delta, t) = V(\delta, t)/Z(t)$ is a Q_Z-martingale. Hence, for a replicating strategy δ_H that replicates the derivative security $H(T)$ we obtain

$$E^{Q_Z}\left(\frac{H(T)}{Z(T)} \,\Big|\, \mathcal{F}_t\right) = E^{Q_Z}\left(\frac{V(\delta_H, T)}{Z(T)} \,\Big|\, \mathcal{F}_t\right) = \frac{V(\delta_H, t)}{Z(t)}, \tag{8}$$

where the last equality follows from the definition of a martingale. Combining the first and last expression yields

$$V(\delta_H, t) = Z(t) E^{Q_Z}\left(\frac{H(T)}{Z(T)} \,\Big|\, \mathcal{F}_t\right). \tag{9}$$

This formula can be used to determine the value at time $t < T$ for any derivative security $H(T)$.

The theorem of the Unique Equivalent Martingale Measure was first discovered by Harrison and Kreps (1979). In their paper they used the value of a riskless money-market account as the numeraire. Later it was recognised that the choice of numeraire is arbitrary. However, for this historic reason, the unique equivalent martingale measure obtained by taking the value of a money-market account as a numeraire, is called "the" equivalent martingale measure, which is a very unfortunate name. In this thesis we will stick to this convention, because it is so widely used.

To illustrate the concepts developed here, we will apply them to the well known Black-Scholes (1973) framework. However, before we do so, we show two results which can be used for explicit calculations in the case of Brownian Motions.

1.3. Girsanov's Theorem and Itô's Lemma

A key result which can be used to explicitly determine equivalent martingale measures in the case of Brownian Motions is *Girsanov's Theorem.* This theorem provides us with a tool to change the drift of a Brownian Motion by changing the probability measure.

Girsanov's Theorem. *For any stochastic process $\kappa(t)$ such that*

$$\int_0^t \kappa(s)^2 \, ds < \infty,$$

with probability one, consider the Radon-Nikodym derivative $\rho(t)$ given by

$$\rho(t) = \exp\Big\{ \int_0^t \kappa(s) \, dW(s) - \tfrac{1}{2} \int_0^t \kappa(s)^2 \, ds \Big\}$$

then under the measure $dQ^* = \rho \, dQ$ the process

$$W^*(t) = W(t) - \int_0^t \kappa(s) \, ds$$

is also a Brownian Motion.

Dothan (1990) provides several generalisations and proofs of this theorem. ■

From the last equation in Girsanov's Theorem follows that

$$dW = dW^* + \kappa(t) \, dt \tag{10}$$

which is a result we will often use.

Another key result from stochastic calculus is known as *Itô's Lemma.* Given a stochastic process x described by a stochastic differential equation, Itô's Lemma allows us to describe the behaviour of stochastic processes derived as functions $f(t, x)$ of the process x.

Itô's Lemma. *Suppose we have a stochastic process x given by the stochastic differential equation $dx = \mu(t, \omega) \, dt + \sigma(t, \omega) \, dW$ and a function $f(t, x)$ of the process x, then f satisfies*

$$\begin{aligned} df = \left(\frac{\partial f(t,x)}{\partial t} + \mu(t,\omega) \frac{\partial f(t,x)}{\partial x} + \tfrac{1}{2} \sigma(t,\omega)^2 \frac{\partial^2 f(t,x)}{\partial x^2} \right) dt \\ + \sigma(t,\omega) \frac{\partial f(t,x)}{\partial x} \, dW, \end{aligned}$$

provided that f is sufficiently differentiable.

For a proof of Itô's Lemma, see Arnold (1992), Sections 5.3 and 5.4. ■

1.4. Application: Black-Scholes Model

Let us now consider the Black and Scholes (1973) option pricing model. Using this familiar setting enables us to illustrate the concepts developed. In the Black-Scholes economy there are two marketed assets: B which is the value of a riskless money-market account with $B(0) = 1$ and a stock S. The prices of the assets are described by the following stochastic differential equations

$$\begin{aligned} dB &= rB\,dt \\ dS &= \mu S\,dt + \sigma S\,dW. \end{aligned} \tag{11}$$

The money-market account is assumed to earn a constant interest rate r, and the stock price is assumed to follow a geometric Brownian Motion with constant drift μ and constant volatility σ.

The value of the money-market account is strictly positive and can serve as a numeraire. Hence, we obtain the relative price $S'(t) = S(t)/B(t)$. From Itô's Lemma we obtain that the relative price process follows

$$dS' = (\mu - r)S'\,dt + \sigma S'\,dW. \tag{12}$$

To identify equivalent martingale measures we can apply Girsanov's Theorem. For $\kappa(t) \equiv -(\mu - r)/\sigma$ we obtain the new measure $dQ^* = \rho\,dQ$ where the process S' follows

$$\begin{aligned} dS' &= (\mu - r)S'\,dt + \sigma S'\left(dW^* - \tfrac{\mu - r}{\sigma}\,dt\right) \\ &= \sigma S'\,dW^* \end{aligned} \tag{13}$$

which is a martingale. For $\sigma \neq 0$ this is the only measure which turns the relative prices into martingales, and the measure Q^* is unique. Therefore, the Black-Scholes economy is arbitrage-free and complete for $\sigma \neq 0$.

Under the measure Q^*, the original price process S follows the process

$$\begin{aligned} dS &= \mu S\,dt + \sigma S\left(dW^* - \tfrac{\mu - r}{\sigma}\,dt\right) \\ &= rS\,dt + \sigma S\,dW^*. \end{aligned} \tag{14}$$

We see that under the equivalent martingale measure the drift μ of the process S is replaced by the interest rate r. The solution to this stochastic differential equation can be expressed as

$$S(t) = S(0)\exp\{(r - \tfrac{1}{2}\sigma^2)t + \sigma W^*(t)\}, \tag{15}$$

where $W^*(t)$ is the value of the Brownian Motion at time t under the equivalent martingale measure. The random variable $W^*(t)$ has a normal distribution with mean 0 and variance t.

A European call option with strike K has at the exercise time T a payoff of $H(T) = \max\{S(T) - K, 0\}$. From (9) follows that the price of the option $\mathbf{C}(0)$ at time 0 is given by $E^*\big(\max\{S(T) - K, 0\}/B(T)\big)$. To evaluate this expectation, we use the explicit solution of $S(T)$ under the equivalent martingale measure given in (15) and we get

$$\begin{aligned} &E^*\big(\max\{S(T) - K, 0\}/B(T)\big) = \\ &\qquad \int_{-\infty}^{\infty} e^{-rT} \max\{S(0)e^{(r-\frac{1}{2}\sigma^2)T+\sigma w} - K, 0\}\frac{e^{-\frac{1}{2}\frac{w^2}{T}}}{\sqrt{2\pi T}}\,dw. \end{aligned} \tag{16}$$

A straightforward calculation will confirm that this integral can be expressed in terms of cumulative normal distribution functions $N(.)$ as follows

$$\mathbf{C}(0) = S(0)N(d) - e^{-rT}KN(d - \sigma\sqrt{T}) \tag{17}$$

with

$$d = \frac{\log\left(\frac{S(0)}{K}\right) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \tag{18}$$

which is the celebrated Black-Scholes option pricing formula.

In the derivation given above, we used the value of a money-market account B as a numeraire. However, this choice is arbitrary. The stock price S is also strictly positive for all t and can also be used as a numeraire. If we choose S as a numeraire, we obtain from Itô's Lemma that the relative price $B' = B/S$ follows

$$dB' = (r - \mu + \sigma^2)B'\,dt - \sigma B'\,dW. \tag{19}$$

If we apply Girsanov's Theorem with $\kappa = (r - \mu)/\sigma + \sigma$, we obtain (for $\sigma \neq 0$) the unique equivalent martingale measure Q_S for which the relative price B' is a martingale. From (9) we obtain

$$\begin{aligned} \mathbf{C}(0) &= S(0)E^{Q_S}\left(\frac{\max\{S(T)-K,0\}}{S(T)}\right) \\ &= S(0)E^{Q_S}\left(\max\{1-K\tfrac{1}{S(T)},0\}\right). \end{aligned} \tag{20}$$

Using Itô's Lemma and Girsanov's Theorem we obtain for the equivalent martingale measure Q_S, that the process $1/S$ follows

$$\begin{aligned} d\frac{1}{S} &= (-\mu+\sigma^2)\frac{1}{S}\,dt - \sigma\frac{1}{S}\left(dW^S + (\tfrac{r-\mu}{\sigma}+\sigma)dt\right) \\ &= -r\frac{1}{S}\,dt - \sigma\frac{1}{S}\,dW^S, \end{aligned} \tag{21}$$

where W^S is a Brownian Motion under Q_S. The explicit solution can be expressed as

$$\frac{1}{S(t)} = \frac{1}{S(0)}\exp\{(-r-\tfrac{1}{2}\sigma^2)t - \sigma W^S(t)\}. \tag{22}$$

Using this explicit form, we can evaluate the expectation (20). It is left to the reader to verify that this also gives the Black-Scholes formula (17).

1.5. Application: Foreign-Exchange Options

The example for the Black-Scholes economy given above is a bit artificial, however a more fruitful application can be found when we consider foreign-exchange (F/X) options. The first valuation formula for F/X-options in a Black-Scholes setting was given by Garman and Kohlhagen (1983). This formula is nowadays widely used by F/X-option traders all over the world.

An interesting aspect of F/X-derivatives is that we can either calculate the value of a derivative in the domestic market or in the foreign market. If the economy is arbitrage-free, both values must be the same, otherwise an "international" arbitrage opportunity would arise.

Consider the following, very simple, international economy. In the domestic market D there is a money-market account B^D, which earns an instantaneous riskless interest rate r^D; in the foreign country F there is also a money-market account B^F with interest rate r^F. Furthermore, the exchange rate X follows a geometric Brownian Motion. The three price processes can be summarised as

$$\begin{cases} dB^F &= r^F B^F \, dt \\ dX &= \mu X \, dt + \sigma X \, dW \\ dB^D &= r^D B^D \, dt \end{cases} \tag{23}$$

From a domestic point of view, there are two marketed assets: the domestic money-market account B^D and the value of the foreign money-market account in domestic terms, given by $B^F X$. From Itô's Lemma we obtain that the process $(B^F X)$ follows

$$d(B^F X) = (r^F + \mu)(B^F X) \, dt + \sigma(B^F X) \, dW. \tag{24}$$

The domestic money-market account can be used as a numeraire, and the relative price process $(B^F X)' = (B^F X)/B^D$ follows the process

$$d(B^F X)' = (r^F - r^D + \mu)(B^F X)' \, dt + \sigma(B^F X)' \, dW. \tag{25}$$

An application of Girsanov's Theorem with $\kappa(t) \equiv -(r^F - r^D + \mu)/\sigma$ will yield the domestic unique equivalent martingale measure Q^D under which the relative price process $(B^F X)'$ is a martingale. Under the domestic measure Q^D, the exchange rate process follows

$$dX = (r^D - r^F) X \, dt + \sigma X \, dW^D, \tag{26}$$

which is the process used in the Garman-Kohlhagen formula.

We can also take the perspective of the foreign market. Here we also have two marketed assets: B^F and (B^D/X). Using B^F as a numeraire, we obtain the relative price process $(B^D/X)' = (B^D/X)/B^F$ which follows the process

$$d\left(\frac{B^D}{X}\right)' = (r^D - r^F - \mu + \sigma^2)\left(\frac{B^D}{X}\right)' dt - \sigma\left(\frac{B^D}{X}\right)' dW. \tag{27}$$

If we apply Girsanov's Theorem with $\kappa(t) \equiv (r^D - r^F - \mu)/\sigma + \sigma$, we obtain the foreign unique equivalent martingale measure Q^F. Under the foreign measure Q^F, the foreign exchange rate $1/X$ follows the process

$$d\left(\frac{1}{X}\right) = (r^F - r^D)\left(\frac{1}{X}\right) dt - \sigma\left(\frac{1}{X}\right) dW^F. \tag{28}$$

This process is exactly the right process for calculating the Garman-Kohlhagen formula in the foreign market. Hence, in this economy a trader in the domestic market and a trader in the foreign market will calculate exactly the same price for a F/X-option.

For more examples of calculating prices of derivatives under domestic and foreign martingale measures, see Reiner (1992). He uses these measures repeatedly to calculate the value of so-called *quanto options*, which are options on foreign assets denominated in the domestic currency.

2. Partial Differential Equations

In this section we give a brief overview of a different methodology for valuing options. By exploiting the fact that for every financial instrument a replicating portfolio can be found and by using no-arbitrage arguments, a partial differential equation can be derived that describes the value of a financial instrument through time.

Given the fact that efficient numerical methods can be used to solve partial differential equations, it is often possible to obtain an accurate approximation of the price of a financial instrument from a partial differential equation in cases where the explicit evaluation of the expectation under the equivalent martingale measure is very difficult. One of the best known examples is probably the pricing of American-style options.

The academic finance literature devotes relatively little attention to the subject of partial differential equations, because it is considered to be more an engineering than an academic problem. However, we believe that no option pricing model can be implemented successfully without a thorough understanding of this subject. The interested reader is referred to the book by Wilmott, Dewynne and Howison (1993), which is completely devoted to the use and solution methods of partial differential equations applied to option pricing theory.

First, we show how in the Black-Scholes economy the partial differential equation can be derived. This is in fact the method employed in Black and Scholes (1973). Then, we explore several analytical and numerical methods to solve partial differential equations, and we establish the equivalence of the equivalent martingale measure methodology and the partial differential equation methodology to calculate the prices of financial instruments.

2.1. Derivation of Black-Scholes Partial Differential Equation

As in Section 1.4 we assume a continuous-time economy with two marketed assets B and S that follow the processes given in (11). We furthermore assume that the value V of a financial instrument is completely determined at every instant t by the asset price $S(t)$. Hence, the value is a function $V(t, S)$. By making this assumption we restrict ourselves to financial instruments whose value does not depend on the history of asset prices until time t. Applying Itô's Lemma gives the following stochastic differential equation for V

$$dV = \left(V_t + \mu S V_S + \tfrac{1}{2}\sigma^2 S^2 V_{SS}\right) dt + \sigma S V_S \, dW, \tag{29}$$

where subscripts denote derivatives.

Suppose we construct a portfolio Π consisting of a position in the instrument V and a short position of Δ assets. The value of the portfolio is then given by $\Pi = V - \Delta S$. Itô's Lemma gives

$$d\Pi = \left(V_t + \mu S(V_S - \Delta) + \tfrac{1}{2}\sigma^2 S^2 V_{SS}\right) dt + \sigma S(V_S - \Delta) \, dW. \tag{30}$$

If we choose $\Delta = V_S$, we see that the dW-term disappears and that the portfolio Π becomes a locally riskless portfolio over the short time interval dt. To avoid arbitrage opportunities, the portfolio Π must earn the same rate of return r as the riskless money-market account B over the time interval dt. Hence, for this choice of Δ we obtain

$$d\Pi = r\Pi \, dt = r(V - V_S S) \, dt. \tag{31}$$

Combining (31) with (30) and dividing by dt leads to

$$V_t + rSV_S + \tfrac{1}{2}\sigma^2 S^2 V_{SS} - rV = 0, \tag{32}$$

which is the Black-Scholes partial differential equation.

Prices of financial instruments can be calculated by solving the partial differential equation with respect to a boundary condition that describes the payoff of the instrument at time T. For example, the price of a European option is given by solving (32) with respect to the boundary condition describing the payoff at time T, namely $V(T, S) = \max\{S(T) - K, 0\}$.

2.2. Fundamental Solutions

Consider functions $V(t,x)$. Partial differential equation of the form

$$V_t + \mu(t,x)V_x + \tfrac{1}{2}\sigma(t,x)^2 V_{xx} - r(t,x)V = 0 \tag{33}$$

like the Black-Scholes partial differential equation (32), have the property that if two functions $V_1(t,x)$ and $V_2(t,x)$ both satisfy the same partial differential equation (disregarding any boundary conditions), then also the linear combination $a_1V_1(t,x)+a_2V_2(t,x)$ satisfies this partial differential equation. It is easily seen that this is true, due to the linearity of the differentiation operator. Even if we have an infinite parameter-family of functions $V(t,x;y)$ that satisfy the same partial differential equation for every y, then the "infinite linear combination" $\int a(y)V(t,x;y)\,dy$ also satisfies this partial differential equation.[6] Suppose now, that we have a (double indexed) parameter-family of functions $V^\delta(t,x;T,y)$ that satisfy a partial differential equation with boundary condition

$$V^\delta(T,x;T,y) = \delta(x-y) \tag{34}$$

for all y and all T. So, every function $V^\delta(t,x;T,y)$ collapses at time T into a Dirac delta-function[7] centred at point y. Consider now a boundary condition described by the function $H(T,x)$ at time T, then the function

$$V(t,x;T) = \int H(T,y)V^\delta(t,x;T,y)\,dy \tag{35}$$

satisfies the partial differential equation and satisfies the boundary condition at $t=T$. The last claim is true since at $t=T$ we have

$$\begin{aligned} V(T,x;T) &= \int H(T,y)V^\delta(T,x;T,y)\,dy \\ &= \int H(T,y)\delta(x-y)\,dy \\ &= H(T,x), \end{aligned} \tag{36}$$

[6] The interchange of differentiation and integration is allowed for suffiently well-behaved functions a and V.

[7] For applications of delta-functions and fundamental solutions, see Griffel (1993) or Williams (1980).

where the last equality follows from the definition of the delta-function.

Hence, for *any* given boundary condition $H(T,x)$, the functions V^δ can be used to construct a solution to the partial differential equation. This is the reason why the functions V^δ are called *fundamental solutions* to this partial differential equation.

If the partial differential equation describes the development of prices through time in an economy (as, e.g. (32) does), we can give an economic interpretation for (35). The Dirac delta-function can be interpreted as the continuous equivalent of the payoff of an Arrow-Debreu security. Because $\delta(x-y) \neq 0$ only for $y = x$ and $\int \delta(x-y)\,dy = 1$, we could say that the delta-function gives a payoff worth 1 in the state of the world $y = x$. Hence, $V^\delta(t,x;T,y)$ can be viewed as the price at time t in state x of an Arrow-Debreu security that has a payoff of 1 at time T in state y. For discrete economies, it is well known that the price of any security with known payoffs at time T can be viewed as a portfolio of Arrow-Debreu securities and can be priced as the payoff-weighted sum over all states of the prices of the Arrow-Debreu securities. It is clear that (35) is the continuous equivalent of this payoff-weighted sum. We will expand on this point in Chapter 4.

2.3. Feynman-Kac Formula

Another method of solving partial differential equations like (32), is to use the *Feynman-Kac formula.*

Feynman-Kac formula. *The partial differential equation*

$$V_t + \mu(t,x)V_x + \tfrac{1}{2}\sigma(t,x)^2 V_{xx} - r(t,x)V = 0$$

with boundary condition $H(T,x)$ has solution

$$V(t,x) = E\left(e^{-\int_t^T r(s,X)\,ds} H(T,X)\right),$$

where the expectation is taken with respect to the process X defined by

$$dX = \mu(t,X)\,dt + \sigma(t,X)\,dW.$$

A proof of a simplified version of the Feynman-Kac formula can be found in Duffie (1988). Proofs of generalised versions of the Feynman-Kac formula can be found in Rogers and Williams (1994). ■

The Black-Scholes partial differential equation (32) can be solved with the Feynman-Kac formula. If we substitute $x = S$, $\mu(t,x) = rS$, $\sigma(t,x) = \sigma S$ and $r(t,x) = r$, we can express the solution with respect to a final payoff function $H(T,x)$ as

$$V(t,S) = E\left(e^{-r(T-t)}H(T,S^*)\right), \tag{37}$$

where E denotes the expectation with respect to the process S^*

$$dS^* = rS^*\,dt + \sigma S^*\,dW. \tag{38}$$

We see that the expectation operator E with respect to the process S^* is exactly the same as the expectation of the discounted payoff under the equivalent martingale measure Q^* if we choose B as numeraire. Hence, calculating the Black-Scholes formula using the partial differential equation leads to exactly the same result (17) which was obtained in Section 1.

2.4. Numerical Methods

Although we have put a lot of emphasis on analytical formulæ up until now, it is often true that prices for options cannot be calculated analytically. An example is, of course, the price of American-style options. Fortunately, there are efficient numerical methods for solving partial differential equations. These numerical methods can be employed in cases where analytical methods fail.

All numerical methods are based on the following method. The solution of the partial differential equation $V(t,S)$ is approximated by values $V_{i,j}$ on a grid with points (t_i, S_j). On the grid the partial derivatives of the solution can be calculated via finite difference approximations. The partial differential equation imposes a relation between the "spatial" derivatives (S-derivatives) and the time-derivative. This relation is then used to propagate the solution from the boundary condition at time T backward to the initial time 0.

Different choices for the grid spacing, and different differencing schemes lead to different algorithms for solving partial differential equations. For detailed derivations of several algorithms, and a discussion of stability and convergence for different algorithms, see Wilmott, Dewynne and Howison (1993). For an introduction to the subject of numerical methods see Hull (1993). In Chapters 5 and 6 of this thesis we provide a derivation of *explicit finite difference* algorithms for interest rate models.

Part 1:

Efficient Methods for Valuing and Managing Interest Rate Derivative Securities

3

Pricing Interest Rate Derivatives

CHAPTER 2 summarises the theory behind the valuation of derivative securities in general. However, in the examples we have given, we made the assumption that it is possible to trade in the underlying values. In the Black-Scholes economy, the underlying value is the asset price in which the agents in the economy can trade.

When we want to consider the valuation of interest rate derivatives, the picture is slightly different. The underlying value of interest rate derivatives is the *spot interest rate*. The spot interest rate is the instantaneous riskless return earned by putting your money for a very short instant dt in a money-market account.

All other interest rates are derived from the spot interest rate. These rates are just different ways of quoting the current prices of discount bonds traded in the market. The prices of discount bonds are determined by the spot interest rate. However, we cannot trade in the spot interest rate itself (nor in any other interest rate), it is only possible to trade in interest rate derivatives like discount bonds.

In this chapter we explain the theory of the pricing of interest rate derivatives, and we point out the implications of the fact that we cannot trade in the spot interest rate. In Section 1 we give a classification of interest rate models. Section 2 is devoted to the valuation of interest rate derivatives using a partial differential equation approach, using the methodology of Vasicek (1977). In Section 3 we explain how interest rate derivatives can be priced via expectations under the equivalent martingale measure, using the methodology of Heath, Jarrow and Morton (1992). In the final section we discuss similarities between the two methodologies.

1. Classification

In this section we provide a classification[1] of different types of interest rate models that have been proposed in the literature and we explain why we focus on one particular type of models.

1.1. Direct/Indirect Approach

One of the simplest methods for valuing interest rate derivatives is to treat the underlying values of interest rate derivatives as if they were traded assets. This approach is widely used for valuing European style interest rate options (caps/floors) and European style bond options. In both cases options are calculated as options on forward prices using the Black (1976) formula. We will call this approach the *direct approach*. Although prices are simple to calculate in the direct approach, it is often the case that these models are internally inconsistent and therefore not arbitrage-free. For example, if one prices both caps/floors and bond options with the Black model both the interest rates and the bond prices are assumed to follow a log-normal distribution, which is clearly inconsistent.

The opposite of the direct approach is the *indirect approach*, which models the behaviour of the spot interest rate, and derives all prices from this process. This ensures that all prices are internally consistent and that the model is arbitrage-free.

1.2. One-Factor/Multi-Factor

Within the class of indirect models, we can provide a further classification. The first classification criterion used for indirect models is the number of sources of uncertainty that drives the evolution of the spot interest rate. Usually a distinction is made between models with only one source of uncertainty called *one-factor models*, and models with more than one source of uncertainty called *multi-factor models*.

The advantage of one-factor models is that they are simple to analyse and understand. The disadvantage of one-factor models is that they

[1] This classification is partially based on the classification given in De Munnik (1992).

severely restrict the types of movement that the term-structure of interest rates can make. Generally, the dynamics of one-factor models display only "monotone" shifts of the term-structure. By monotone we mean that all points on the curve move in the same direction.

Multi-factor models display a much richer behaviour of the term-structure, which is more realistic. However, the price that has to be paid is that multi-factor models are much more difficult to analyse and are computationally much more cumbersome.

1.3. Endogenous/Exogenous Term-Structure of Interest Rates

Indirect models can be classified in another way. The first approach (and historically the oldest) is to describe the spot interest-rate by a stochastic differential equation with constant parameters. This leads to models which have an *endogenous term-structure of interest rates*. The best known examples are the one-factor models proposed by Vasicek (1977) and Cox, Ingersoll and Ross (1985). Due to the fact that the term-structure of interest rates is endogenously given in these models, it is not possible to fit the observed term-structure of interest rates (and hence the observed bond- or swap-prices) exactly. Hence, these models are not consistent with the prices traded in the market.

This shortcoming has inspired the second type of approach, which takes the initial term-structure of interest rates as given, and describes the dynamics of the complete term-structure. These are models with an *exogenous term-structure of interest rates*. We will also call this type of models *yield-curve models*. The first yield-curve model was the one-factor model proposed by Ho and Lee (1986). Other well-known models are the one- and multi-factor yield-curve models of Hull and White (1990a, 1994).

For the remainder of this thesis we will restrict ourselves to one-factor yield-curve models, which are indirect one-factor models with an exogenous term-structure of interest rates. The reason for focussing on this type of models is that they are arbitrage-free models which can fit the observed bond prices, and that they are computationally efficient.

2. Vasicek Methodology

In this section we show how interest rate derivatives can be valued using partial differential equations. The derivation of the partial differential equation is based on Vasicek (1977).

2.1. Spot Interest Rate

Indirect one-factor models describe the evolution of the spot interest rate with one source of uncertainty. We can write down the following general stochastic differential equation for the spot interest rate r

$$dr = \mu(t,r)\,dt + \sigma(t,r)\,dW. \tag{1}$$

In this general form the functions $\mu(t,r)$ and $\sigma(t,r)$ are left unspecified. Different choices for the functions μ and σ give rise to different models.

Given this stochastic process for the spot interest rate r we can proceed to derive a partial differential equation by constructing a locally riskless portfolio.

2.2. Partial Differential Equation

In analogy to Chapter 2 we want to consider the value V of financial instruments whose value is determined at time t by the value of the spot interest rate $r(t)$. The value of a financial instrument V is a function $V(t,r)$. From Itô's Lemma we obtain

$$dV = M(t,r)\,dt + \Sigma(t,r)\,dW, \tag{2}$$

with

$$\begin{aligned} M(t,r) &= V_t + \mu(t,r)V_r + \tfrac{1}{2}\sigma(t,r)^2 V_{rr} \\ \Sigma(t,r) &= \sigma(t,r)V_r \end{aligned} \tag{3}$$

Let us now attempt to construct a locally riskless portfolio Π. We would like to take a position in the instrument V with a short position in the spot interest rate r. Unfortunately, the spot interest rate r is not a traded asset, so this is impossible. The best thing we can do, is to hedge a derivative V_1 with another interest rate derivative V_2 to obtain the portfolio

$$\Pi = V_1(t,r) - \Delta V_2(t,r), \tag{4}$$

where V_1 and V_2 follow processes similar to (2), with coefficients M_1, Σ_1 and M_2, Σ_2, respectively.

The portfolio Π is a linear combination of the stochastic processes V_1 and V_2. Hence, we obtain

$$d\Pi = \big(M_1(t,r) - \Delta M_2(t,r)\big)dt + \big(\Sigma_1(t,r) - \Delta\Sigma_2(t,r)\big)dW. \tag{5}$$

For the choice $\Delta = \Sigma_1/\Sigma_2$ the dW disappears and the portfolio becomes locally riskless for the time interval dt. To avoid arbitrage opportunities, the portfolio must earn in this case the locally riskless return, which is the spot interest rate r. Hence, for this choice of Δ we obtain for the time period dt

$$d\Pi = r\Pi\, dt$$
$$\Downarrow$$
$$\left(M_1(t,r) - \frac{\Sigma_1(t,r)}{\Sigma_2(t,r)} M_2(t,r)\right) dt = r\left(V_1(t,r) - \frac{\Sigma_1(t,r)}{\Sigma_2(t,r)} V_2(t,r)\right) dt. \tag{6}$$

Dividing by dt and rearranging terms leads to

$$\frac{M_1(t,r) - rV_1(t,r)}{\Sigma_1(t,r)} = \frac{M_2(t,r) - rV_2(t,r)}{\Sigma_2(t,r)}. \tag{7}$$

This equality must hold for any pair of derivatives V_1 and V_2, which is only possible if the value of the ratio $(M - rV)/\Sigma$ is a function of t and r only. Let $\lambda(t,r)$ denote the common value of the ratio, which is known as the *market price of risk*. Hence, any interest rate derivative V must satisfy

$$\frac{M(t,r) - rV(t,r)}{\Sigma(t,r)} = \lambda(t,r), \tag{8}$$

where M and Σ are defined as in (3). Substituting these definitions into (8) and rearranging terms yields

$$V_t + \big(\mu(t,r) - \lambda(t,r)\sigma(t,r)\big)V_r + \tfrac{1}{2}\sigma(t,r)^2 V_{rr} - rV = 0 \tag{9}$$

which is the partial differential equation that describes the prices of securities in one-factor yield-curve models.

2.3. Calculating Prices

Armed with the partial differential equation (9), the price of an interest rate derivative V with a payoff $V(T,r)$ at time T can be calculated with the help of the Feynman-Kac formula. The price $V(t,r)$ is the solution to the partial differential equation subject to the boundary condition $V(T,r)$ and can be expressed as

$$V(t,r) = E\left(e^{-\int_t^T r^*(s)\, ds} V(T, r^*)\right), \tag{10}$$

where the expectation is taken with respect to the process r^*

$$dr^* = \theta(t, r^*)\, dt + \sigma(t, r^*)\, dW, \tag{11}$$

where

$$\theta(t, r^*) = \mu(t, r^*) - \lambda(t, r^*)\sigma(t, r^*). \tag{12}$$

The process r^* used to calculate the expectation is different from the process r defined in (1). The dW term of both processes is the same, but the drift term of r^* is corrected by a factor involving the market price of risk $\lambda(t, r)$.

Before prices of derivatives can be calculated, the model has to be fitted to the initial term-structure of interest rates. The initial term-structure of interest rates is described by the prices of all discount bonds $P(0, T)$ at time $t = 0$. The payoff of a discount bond at maturity T is 1 in all states of the world. Hence, using the Feynman-Kac formula (10) we can express the discount bond prices in the functions θ and σ. Given a choice for σ, we can solve for θ from the initial term-structure of interest rates. When we have determined θ, we have actually estimated the drift μ and the market price of risk λ simultaneously from the initial term-structure of interest rates. The valuation formula (10) is then no longer (explicitly) dependent on the market price of risk.

2.4. Example: Ho-Lee Model

To illustrate the procedure outlined above we provide a simple example. If we assume that $\sigma(t, r)$ is a constant σ and that $\mu(t, r)$ is a function $\mu(t)$ of time only, we obtain the continuous time limit of the Ho and Lee (1986) model. For these choices of μ and σ the partial differential equation (9) reduces to

$$V_t + \big(\mu(t) - \lambda(t, r)\sigma\big) V_r + \tfrac{1}{2}\sigma^2 V_{rr} - rV = 0. \tag{13}$$

If we make the additional assumption that the market price of risk is a function $\lambda(t)$ of time only, the drift term is a function of time only which can be denoted by $\theta(t)$.

Using the Feynman-Kac formula, the prices of interest rate derivatives can be expressed as

$$V(t, r) = E\left(e^{-\int_t^T r^*(s)\, ds} V(T, r^*)\right), \tag{14}$$

where the expectation is taken with respect to the process r^*

$$dr^* = \theta(t)\,dt + \sigma\,dW. \tag{15}$$

To fit this model to the initial term-structure of interest rates, we have to calculate the prices of discount bonds in terms of $\theta(t)$. The payoff of a discount bond at maturity is equal to 1, hence we have $V(T, r^*) \equiv 1$ and the price of a discount bond is given by

$$P(0,T) = E\left(e^{-y(T)}\right), \tag{16}$$

where the random variable y is defined as

$$y(t) = \int_0^t r^*(s)\,ds. \tag{17}$$

Substituting the solution of the stochastic differential equation (15) for r^* into the definition of $y(t)$ yields

$$y(t) = \int_0^t r_0\,ds + \int_0^t \int_0^s \theta(u)\,du\,ds + \int_0^t \int_0^s \sigma\,dW(u)\,ds. \tag{18}$$

By interchanging the order of integration[2] and simplifying we obtain

$$y(t) = r_0 t + \int_0^t \theta(u)(t-u)\,du + \int_0^t \sigma(t-u)\,dW(u). \tag{19}$$

Hence, the process $y(t)$ has a normal distribution with mean

$$m(t) = r_0 t + \int_0^t \theta(s)(t-s)\,ds \tag{20}$$

and variance

$$v(t) = \int_0^t \sigma^2(t-s)^2\,ds = \tfrac{1}{3}\sigma^2 t^3. \tag{21}$$

[2] For a proof of Fubini's Theorem for stochastic integrals, see the Appendix of Heath, Jarrow and Morton (1992).

From this follows that the expectation in the Feynman-Kac formula (16) can be evaluated as

$$\begin{aligned} P(0,T) &= \exp\{-m(T) + \tfrac{1}{2}v(T)\} \\ &= \exp\Big\{-r_0 T - \int_0^T \theta(s)(T-s)\,ds + \tfrac{1}{6}\sigma^2 T^3\Big\}. \end{aligned} \tag{22}$$

The Ho-Lee model can be fitted to the initial term-structure of interest rates by solving for $\theta(t)$. Taking logarithms and differentiating twice with respect to T yields

$$\theta(T) = -\frac{\partial^2}{\partial T^2}\log P(0,T) + \sigma^2 T. \tag{23}$$

3. Heath-Jarrow-Morton Methodology

In this section we show how prices of interest rate derivatives can be calculated using equivalent martingale measures, which is the methodology of Heath, Jarrow and Morton (1992). We will only derive the equivalent martingale measure for one-factor interest rate models, this will allow us to explain the essence of the Heath-Jarrow-Morton (HJM) methodology.

3.1. Forward Rates

The marketed assets which can be traded are discount bonds with different maturities. The price of a discount bond at time t with maturity T is denoted by $P(t,T)$. In their setup HJM choose not to model discount bond prices directly, but to model the prices of *forward rates* $f(t,T)$. The forward rate is given by

$$f(t,T) = \frac{-\partial \log P(t,T)}{\partial T}, \tag{24}$$

it is the instantaneous interest rate one can contract for at time t to invest in the money-market account at time T. It is easy to see that the spot interest rate $r(t)$ is equal to $f(t,t)$.

HJM assume that the forward rates satisfy the following equation

$$f(t,T) - f(0,T) = \int_0^t \alpha(s,T,\omega)\,ds + \int_0^t \sigma(s,T,\omega)\,dW(s), \tag{25}$$

where ω denotes the state of the world. Equation (25) is the integral form of the stochastic differential equation

$$df(t,T) = \alpha(t,T,\omega)\,dt + \sigma(t,T,\omega)\,dW, \tag{26}$$

however, the integral form (25) of the equation is more precise. The stochastic process for the forward rates defined above is very general. The functions α and σ are allowed to depend on the maturity T of the forward rate and are allowed to depend on the state of the world ω.

The spot interest rate $r(t)$ is equal to $f(t,t)$, hence we get from (25)

$$r(t) = f(0,t) + \int_0^t \alpha(s,t,\omega)\,ds + \int_0^t \sigma(s,t,\omega)\,dW(s). \tag{27}$$

This stochastic process for the spot rate r is much more general than the process (1) proposed in the previous section. For the appropriate choices for α and σ it is (in principle) possible to reduce (27) to the form (1).

Using (24) we can express the discount bond prices in terms of the forward rates as

$$\log P(t,T) = -\int_t^T f(t,s)\,ds. \tag{28}$$

Substituting (25) into this equation and by interchanging the order of integration and simplifying, HJM obtain the following process for the discount bond prices (suppressing the notational dependence on ω)

$$dP(t,T) = b(t,T)P(t,T)\,dt + a(t,T)P(t,T)\,dW, \tag{29}$$

where

$$\begin{aligned} a(t,T,\omega) &= -\int_t^T \sigma(t,s,\omega)\,ds \\ b(t,T,\omega) &= r(t) - \int_t^T \alpha(t,s,\omega)\,ds + \tfrac{1}{2}a(t,T,\omega)^2. \end{aligned} \tag{30}$$

3.2. Equivalent Martingale Measure

Having specified the stochastic process followed by the discount bonds $P(t,T)$ which are the marketed assets, we want to establish the existence of an equivalent martingale measure to ensure that no arbitrage opportunities can exist in the economy.

Suppose we keep reinvesting money in the money-market account. Every instant dt the money market account earns the riskless spot interest rate and the value $B(t)$ of the money-market account is given by $dB = rBdt$. If we solve this ordinary differential equation we get

$$B(t) = \exp\Big\{\int_0^t r(s)\,ds\Big\}. \tag{31}$$

As in the Black-Scholes economy of Chapter 2, the value of the money-market account is strictly positive and can be used as a numeraire. Hence, in the HJM economy we obtain the relative prices $P'(t,T) = P(t,T)/B(t)$. Itô's Lemma yields

$$dP'(t,T) = \big(b(t,T) - r(t)\big)P'(t,T)\,dt + a(t,T)P'(t,T)\,dW. \tag{32}$$

The HJM economy will be arbitrage-free if we can find a unique equivalent probability measure such that the relative prices P' of the discount bonds become martingales.

Suppose we consider the discount bond with maturity T_1. If we apply Girsanov's Theorem with $\kappa(t,T_1) = -\big(b(t,T_1) - r(t)\big)/a(t,T_1)$, we obtain under the new measure $dQ^*_{T_1} = \rho(t,T_1)\,dQ$ that the process $P'(t,T_1)$ is a martingale. This change of measure depends on the maturity of the discount bond T_1 and will only make this particular discount bond a martingale.

However, we want to find an equivalent martingale measure that changes all marketed assets, that is all discount bonds, to martingales. This is only possible if the ratio $\big(b(t,T,\omega) - r(t)\big)/a(t,T,\omega)$ is independent of T. Let $\lambda(t,\omega)$ denote the common value of this ratio, if we apply then Girsanov's Theorem with $\kappa(t,\omega) = -\lambda(t,\omega)$ we get that all discount bonds $P'(t,T)$ are martingales under the equivalent martingale measure $dQ^* = \rho\,dQ$.

Since the prices of all discount bonds are dependent on the spot interest rate r, the drift term $b(t,T,\omega)$ cannot be specified arbitrarily. A unique equivalent martingale measure can only be found if the drift term is of the form

$$b(t,T,\omega) - r(t) = \lambda(t,\omega)a(t,T,\omega). \tag{33}$$

Substituting the definitions for a and b given in (30), and differentiating with respect to T we find that the drift terms of the forward rate processes $\alpha(t,T,\omega)$ are restricted to

$$\alpha(t,T,\omega) = \sigma(t,T,\omega)\Big(\int_t^T \sigma(t,s,\omega)\,ds + \lambda(t,\omega)\Big). \tag{34}$$

3.3. Calculating Prices

Now that we have determined under which conditions an equivalent martingale measure exists in the HJM model, we can calculate the prices of interest rate derivatives. In Chapter 2 we derived the result that under the equivalent martingale measure the relative prices $V(t,r)/B(t)$ are martingales. In the HJM economy we get that the price of a financial instrument with a payoff $H(T,r)$ at time T is given by

$$V(t,r) = E^*\Big(e^{-\int_t^T r(s)\,ds} H(T,r) \,\Big|\, \mathcal{F}_t\Big) \tag{35}$$

where the expectation E^* is taken with respect to the equivalent martingale measure Q^*. From Girsanov's Theorem and using the restriction on α given in (34) we obtain that under the equivalent martingale measure the process r follows

$$r(t) = f(0,t) + \int_0^t \sigma(s,t,\omega) \int_s^t \sigma(s,u,\omega)\,du\,ds + \int_0^t \sigma(s,t,\omega)\,dW^*(s). \tag{36}$$

It is clear that for a given initial term-structure of interest rates and for a given choice of the function $\sigma(t,T,\omega)$ the spot rate process under the equivalent martingale measure is completely determined.

3.4. Example: Ho-Lee Model

To illustrate the HJM methodology, we turn again to the continuous-time Ho-Lee model. If we set the function $\sigma(t,T,\omega)$ to a constant σ, and if we make the assumption that the market price of risk is a function $\lambda(t)$ of time only, we obtain from (34) that the drift terms of the forward rates are restricted to

$$\alpha(t,T) = \sigma\big(\sigma(T-t) + \lambda(t)\big). \tag{37}$$

Hence, under the equivalent martingale measure, the spot interest rate follows the process

$$r(t) = f(0,t) + \tfrac{1}{2}\sigma^2 t^2 + \sigma W^*(t), \tag{38}$$

which can also be written in differential form as

$$dr = \left(-\frac{\partial^2}{\partial t^2}\log P(0,t) + \sigma^2 t\right) dt + \sigma\, dW^*, \tag{39}$$

where we have used the definition of the forward rates given in (24).

4. Conclusion

If we compare the process r^* from the Feynman-Kac formula with the process r under the equivalent martingale measure we see that they look very different. This has two reasons. First, the process for the forward rates defined by HJM is much more general than the forward rates induced by the process (1). In general, HJM processes are path dependent, which means that the stochastic process at time t depends on the path followed by the forward rates up until time t. Second, the drift term of the process r^* depends implicitly on the forward rate, because we still have to fit it to the initial term-structure of interest rates. The drift term of the process r under the equivalent martingale measure depends explicitly on the initial term-structure of interest rates.

Although the HJM specification is very general, it is not easy to specify for which choices of the function $\sigma(t,T,\omega)$ we obtain the path-independent models of the Vasicek methodology. For some models, like the Ho-Lee model, the choice for σ is known. In these cases one can show that the Vasicek and the HJM methodology are exactly the same. Establishing the general equivalence of both methodologies remains a topic of research.

4

Fundamental Solutions and the Forward-Risk-Adjusted Measure

PRICES of interest rate derivatives can be calculated as the expected value of the discounted payoff, this was explained in the previous chapter. However, interest rates play a double role in interest rate models: they determine the amount of discounting, and they determine the payoff of the security. This implies that the discounting term and the payoff term are two correlated stochastic variables, which makes the evaluation of the expectation quite difficult.

As was shown independently by Jamshidian (1991) and by Geman et al. (1995), one can use the T-maturity discount bond as a numeraire with its associated unique equivalent martingale measure. Under this new measure, which was named the *T-forward-risk-adjusted measure* by Jamshidian, prices of interest rate derivatives can be calculated as the discounted expected value of the payoff, which makes the calculation much simpler. However, explicitly determining this new measure can be complicated.

In this chapter we provide an alternative method to determine the T-forward-risk-adjusted measure for interest rate models. We do so by showing that the fundamental solutions to the pricing partial differential equation can be interpreted as the discounted probability density functions associated with the T-forward-risk-adjusted measure. A method to obtain fundamental solutions from the partial differential equation using Fourier transforms is introduced.

We define the class of normal models. These are interest rate models where the spot interest rate is a deterministic function of an underlying normally distributed stochastic process that drives the economy. We show that the models with the richest analytical structure belong to the class of normal models. These models with a rich analytical structure have also normally distributed fundamental solutions. Using the methods introduced in this chapter we derive an important theoretical result. We prove that within the class of normal models only the set of models where the spot interest rate is either a linear or a quadratic function of the underlying process has normally distributed fundamental solutions.

The rest of this chapter is organised as follows. In Section 1 we show how prices of interest rate derivatives can be calculated in a simpler fashion by changing to the T-forward-risk-adjusted measure. In Section 2 we show how the fundamental solutions to the partial differential equation can be interpreted as the probability density functions under the new measure. A method for obtaining fundamental solutions is explained in Section 3 and an example is given in Section 4 that illustrates the concepts developed so far. Finally, in Section 5, we prove the theoretical result that within the class of normal models only the set of models where the spot interest rate is either a linear or a quadratic function of the underlying process has normally distributed fundamental solutions.

1. Forward-Risk-Adjusted Measure

In Chapter 3 we derived the result that prices of interest rate derivatives can be calculated by solving the appropriate partial differential equation (the Vasicek methodology) or by taking the expectation with respect to the equivalent martingale measure (the HJM methodology). The price $V(t,r)$ of an interest rate derivative can be expressed as

$$V(t,r) = E^*\Big(e^{-\int_t^T r(s)\,ds} H(T, r(T)) \,\Big|\, \mathcal{F}_t\Big), \tag{1}$$

where the expectation is taken with respect to the equivalent martingale measure $Q^*$and $H(T,r)$ denotes the payoff of the derivative at time T.

When we determined the equivalent martingale measure in the HJM economy, we used the value of the money-market account $B(t)$ as a numeraire. However, this is not the only choice we could have made. As was argued in Chapter 2, it is possible to use any financial instrument with a strictly positive price (and no intermediate payouts) as a numeraire. The following theorem shows that if an equivalent martingale measure can be found with respect to one numeraire, then an equivalent martingale measure can be found with respect to any numeraire.

Change of Numeraire Theorem. *Let Q^* be the equivalent martingale measure with respect to a numeraire $B(t)$. Let $X(t)$ be an arbitrary numeraire. Then there exists a probability measure Q_X such that:*

i for any asset $S(t)$ we have $S(t)/X(t)$ is a martingale with respect to Q_X;

ii the Radon-Nikodym derivative that changes the equivalent martingale measure Q^* into Q_X is given by

$$\frac{dQ_X}{dQ^*} = \frac{X(T)}{X(0)B(T)}.$$

A proof can be found in Geman et al. (1995). ■

The Change of Numeraire Theorem is very powerful in the context of valuing interest rate derivatives. Instead of using the value of the money-market account $B(t)$ as a numeraire, the prices of discount bonds $P(t,T)$ can also be used as a numeraire. A very convenient choice is to use the discount bond with maturity T as a numeraire for derivatives which have a payoff $H(T,r(T))$ at time T. If we denote the probability measure associated to the numeraire $P(t,T)$ by Q_T we can apply the Change of Numeraire Theorem as follows. Under the measure Q_T the prices $V(t,r)/P(t,T)$ are martingales for $t < T$. Hence, applying the definition of a martingale we obtain

$$E^{Q_T}\big(V(T,r(T))/P(T,T) \bigm| \mathcal{F}_t\big) = V(t,r)/P(t,T). \tag{2}$$

However, at time T the price of the discount bond $P(T,T) \equiv 1$ and the price of the instrument V is given by its payoff $V(T,r(T)) = H(T,r(T))$. So, the equation reduces to

$$V(t,r) = P(t,T)E^{Q_T}\big(H(T,r(T)) \bigm| \mathcal{F}_t\big). \tag{3}$$

If we compare this expression with (1), we see that we have managed to express the expectation of the discounted payoff as a discounted expectation of the payoff. We have eliminated the problem of the correlation between the discounting term and the payoff term.

The measure Q_T has another very interesting property, which actually gave it the name T-forward-risk-adjusted measure. Under the T-forward-risk-adjusted measure, the forward rate $f(t,T)$ is equal to the expected value of the spot interest rate at time T. The following argument shows why this is true. A discount bond $P(t,T)$ has a payoff of 1 at time T. Using (1), the price of the discount bond can be expressed as

$$P(t,T) = E^*\Big(e^{-\int_t^T r(s)\,ds} \cdot 1 \;\Big|\; \mathcal{F}_t\Big). \tag{4}$$

Taking derivatives with respect to T yields

$$\begin{aligned} -\frac{\partial}{\partial T} P(t,T) &= E^*\Big(e^{-\int_t^T r(s)\,ds} r(T) \;\Big|\; \mathcal{F}_t\Big) \\ &= P(t,T) E^{Q_T}\big(r(T) \;\big|\; \mathcal{F}_t\big), \end{aligned} \tag{5}$$

where we have used (3) in the last step. Using the definition of the forward rates $f(t,T) = -\partial/\partial T \log P(t,T)$ we can simplify this expression to

$$f(t,T) = E^{Q_T}\big(r(T) \;\big|\; \mathcal{F}_t\big), \tag{6}$$

which is the desired result.

Although we have eliminated in expression (3) the problem of the correlation between the discounting term and the payoff term, we have introduced the problem of having to determine the distribution of r under the probability measure Q_T. In some cases it is possible to determine the probability measure directly from the Radon-Nikodym derivative, however, in general this can be complicated. In the next section we explore a different way of determining the distribution of r under the measure Q_T using the concept of fundamental solutions of the partial differential equation.

2. Fundamental Solutions

If we use the Vasicek methodology of Chapter 3, then the prices of interest rate derivatives are described by a partial differential equation of the form

$$V_t + \mu(t,r)V_r + \tfrac{1}{2}\sigma(t,r)^2 V_{rr} - rV = 0. \tag{7}$$

In Chapter 2 we showed that, with the help of fundamental solutions $V^\delta(t,r;T,z)$, solutions to a partial differential equation can be expressed in terms of the boundary condition $H(T,r)$ as follows

$$V(t,r;T) = \int H(T,z) V^\delta(t,r;T,z)\,dz. \tag{8}$$

A discount bond P is a security that gives a payoff of 1 in all states of the world at maturity T. Hence, the price $P(t,T,y)$ at time t of a discount bond is $P(t,T,r) = \int V^{\delta}(t,r;T,z)\,dz$. If the economy is arbitrage-free, prices of Arrow-Debreu securities cannot become negative and prices of discount bonds are finite. These observations lead to another interpretation for (8). The function

$$p^{\delta}(t,r;T,z) = \frac{V^{\delta}(t,r;T,z)}{P(t,T,r)} \tag{9}$$

is non-negative and (by construction) integrates out to 1 with respect to z. Any function that is non-negative and integrates out to 1 can be interpreted as a probability density function, and (8) can be written as

$$\begin{aligned} V(t,r;T) &= P(t,T,r)\int H(T,z)p^{\delta}(t,r;T,z)\,dz \\ &= P(t,T,r)\widetilde{E}^{T}\big(H(T,z) \mid \mathcal{F}_t\big), \end{aligned} \tag{10}$$

where $\widetilde{E}^{T}(.\mid \mathcal{F}_t)$ denotes the expectation operator, conditional on the information available at time t, with respect to the density $p^{\delta}(t,r;T,.\,)$.

If we combine this result with (1) and (3), where z represents $r(T)$, we get that

$$V\big(t,r;T\big) = E^{*}\big(e^{-\int_t^T r(s)\,ds}H(T,z) \mid \mathcal{F}_t\big) = P(t,T,r)\widetilde{E}^{T}\big(H(T,z) \mid \mathcal{F}_t\big). \tag{11}$$

We have expressed the expectation E^{*} of the discounted payoff, as the discounted expectation $\widetilde{E}^{T}$ of the payoff, which is exactly the same result as in Section 1. Since this is true for any payoff H, the probability density functions $p^{\delta}(t,r;T,z)$ must be equal to the transition density functions $p^{Q_T}(t,r;T,z)$ of the process r under the T-forward-risk-adjusted measure Q_T.

3. Obtaining Fundamental Solutions

If we want to find fundamental solutions for the partial differential equation (7), we seek functions $V^{\delta}(t,r;T,z)$ that satisfy the partial differential equation, and collapse into a delta-function at $t = T$ for all T and all z. In order to facilitate the calculations, we will not try to solve for the fundamental solutions directly, but we will solve the partial differential equation for the *Fourier transform*[1] of the fundamental solutions, and obtain the fundamental solutions by inverting the Fourier transform.

Consider the function $\widetilde{V}(t, r; T, \psi)$ defined as

$$\widetilde{V}(t, r; T, \psi) = \int_{-\infty}^{\infty} e^{i\psi z} V^{\delta}(t, r; T, z)\, dz. \tag{12}$$

This function is the Fourier transform in the variable z of V^{δ}, where i is the imaginary number, for which $i^2 = -1$. The function $\widetilde{V}$ satisfies the same partial differential equation as V^{δ}, but the boundary condition is given by

$$\widetilde{V}(T, r; T, \psi) = \int_{-\infty}^{\infty} e^{i\psi z} \delta(r - z)\, dz = e^{i\psi r}, \tag{13}$$

where the last equality follows from the definition of the delta-function. We see that the boundary condition for $\widetilde{V}$ is a very simple function.

The fundamental solutions V^{δ} can be obtained from $\widetilde{V}$ by inverting the Fourier transform. This inversion is often simple, if we use the following property. Using (9) we can write the fundamental solution as the product of the discount bond price times a probability density function $V^{\delta}(t, r; T, z) = P(t, T, y)p^{\delta}(t, r; T, z)$. Hence, we can write $\widetilde{V}$ as

$$\widetilde{V}(t, r; T, \psi) = \int_{-\infty}^{\infty} e^{i\psi z} P(t, T, y) p^{\delta}(t, r; T, z)\, dz = P(t, T, y)\tilde{p}(t, r; T, \psi), \tag{14}$$

where $\tilde{p}$ is the Fourier transform of the density p^{δ}. However, the Fourier transform of a probability density function is the same as the *characteristic function*[2] of that density. For $\psi = 0$, the boundary condition of $\widetilde{V}$ reduces to the boundary condition for a discount bond, and we get

$$\begin{aligned} P(t, T, r) &= \widetilde{V}(t, r; T, 0) \\ \tilde{p}(t, r; T, \psi) &= \frac{\widetilde{V}(t, r; T, \psi)}{\widetilde{V}(t, r; T, 0)} = \frac{\widetilde{V}(t, r; T, \psi)}{P(t, T, r)}. \end{aligned} \tag{15}$$

[1] For applications of the Fourier transform, and for conditions under which the Fourier transform exists, see Duffie (1994).

[2] See, e.g. Lukacs (1970).

So, we see that both the price of a discount bond P, and the characteristic function $\tilde{p}$ of the density p^δ, can be obtained from $\widetilde{V}$. Characteristic functions can be looked up in standard tables, and the inversion is then simple. Even if the characteristic function cannot be inverted, useful information about the distribution, like the moments or approximating distributions, can be obtained from the characteristic function. Once the discount bond price P and the densities p^δ have been obtained, prices for interest rate derivatives can be calculated by using (10).

4. Example: Ho-Lee Model

In this section we provide an example to illustrate the concepts developed so far. We consider the continuous-time limit of the Ho-Lee model we have already encountered in Chapter 3. For this model we can explicitly determine the T-forward-risk-adjusted measure both from the Radon-Nikodym derivative and from the fundamental solutions of the partial differential equation. We show that both derivations are consistent.

Other applications of the concepts developed in this chapter can be found in Chapter 5, where we analyse the Hull-White (1994) model, and in Chapter 6, where we analyse the one-factor squared Gaussian model.

4.1. Radon-Nikodym Derivative

In Chapter 3 we showed that, if we use the value of the money-market account $B(t)$ as a numeraire, under the equivalent martingale measure Q^* the spot interest rate r follows the process

$$dr = \theta^*(t)\,dt + \sigma\,dW^*, \tag{16}$$

where $\theta^*(t)$ is given by

$$\theta^*(t) = -\frac{\partial^2}{\partial t^2}\log P(0,t) + \sigma^2 t. \tag{17}$$

The prices at time $t = 0$ of interest rate derivatives with payoff $H(r(T))$ at maturity T can be calculated as

$$V_0 = E^*\left(e^{-\int_0^T r(s)\,ds} H(r(T))\right). \tag{18}$$

Instead of using the value of the money-market account $B(t)$, we can also use the value of the T-maturity discount bond $P(t,T)$ as a numeraire. In Section 2 we explained that prices can also be calculated as

$$V_0 = P(0,T)E^{Q_T}\big(H(r(T))\big). \tag{19}$$

The Change of Numeraire Theorem provides us with the appropriate Radon-Nikodym derivative to change from Q^* to Q_T

$$\frac{dQ_T}{dQ^*} = \frac{P(T,T)}{P(0,T)B(T)} = \frac{e^{-\int_0^T r(s)\,ds}}{P(0,T)} \tag{20}$$

If we substitute the solution of the stochastic differential equation (16) into this equation and interchange the order of integration we obtain for dQ_T/dQ^*

$$\exp\Big\{-\log P(0,T) - \int_0^T \left(-\tfrac{\partial^2 \log P(0,s)}{\partial s^2} + \sigma^2 s\right)(T-s)\,ds \\ - \int_0^T \sigma(T-s)\,dW^*(s)\Big\}. \tag{21}$$

If we work out the integrals, we see that the Radon-Nikodym derivative can be simplified to

$$\frac{dQ_T}{dQ^*} = \exp\Big\{-\tfrac{1}{6}\sigma^2 T^3 - \int_0^T \sigma(T-s)\,dW^*(s)\Big\}. \tag{22}$$

From Girsanov's Theorem follows that a Radon-Nikodym derivative of this form can be obtained by setting $\kappa(t) = -\sigma(T-t)$. Hence, under the measure Q_T the process

$$W^T(t) = W^*(t) - \int_0^t (-\sigma)(T-s)\,ds \tag{23}$$

is also a Brownian Motion. From this we obtain that for $t < T$ the spot interest rate r follows the process

$$dr = \big(\theta^*(t) - \sigma^2(T-t)\big)\,dt + \sigma\,dW^T \tag{24}$$

under the probability measure Q_T.

4.2. Fundamental Solutions

In Chapter 3 we derived that prices of interest derivatives in the Ho-Lee model follow the partial differential equation

$$V_t + \theta^*(t)V_r + \tfrac{1}{2}\sigma^2 V_{rr} - rV = 0. \tag{25}$$

To find the function $\widetilde{V}(t, r; T, \psi)$ for the Ho-Lee model, we have to solve (25) with respect to the boundary condition $\widetilde{V}(T, r; T, \psi) = e^{i\psi r}$. As a solution we try

$$\widetilde{V}(t, r; T, \psi) = \exp\{A(t; T\psi) + B(t; T, \psi)r\}. \tag{26}$$

Substituting this functional form in (25) and collecting terms yields

$$r(B_t - 1) + A_t + \theta^*(t)B + \tfrac{1}{2}\sigma^2 B^2 = 0, \tag{27}$$

which is solved if A and B solve the following system of ordinary differential equations

$$\begin{cases} B_t - 1 = 0 \\ A_t + \theta^*(t)B + \tfrac{1}{2}\sigma^2 B^2 = 0 \end{cases} \tag{28}$$

subject to the boundary conditions $A(T; T, \psi) = 0$ and $B(T; T, \psi) = i\psi$. The solutions for A and B are given by

$$\begin{aligned} B(t; T, \psi) &= i\psi - (T - t) \\ A(t; T, \psi) &= \tfrac{1}{6}\sigma^2 (T - t)^3 - \int_t^T \theta^*(s)(T - s)\,ds \\ &\quad + i\psi\Big(\int_t^T \theta^*(s) - \sigma^2 (T - s)\,ds\Big) - \tfrac{1}{2}\psi^2\sigma^2 (T - t). \end{aligned} \tag{29}$$

If we substitute these expressions for A and B into (26) we obtain an explicit expression for the function $\widetilde{V}(t, r; T, \psi)$.

In Section 3 we showed that the price of a discount bond can be obtained by setting $\psi = 0$ in $\widetilde{V}$, this leads to

$$P(t, T, r) = \exp\Big\{-r(T - t) + \tfrac{1}{6}\sigma^2 (T - t)^3 - \int_t^T \theta^*(s)(T - s)\,ds\Big\}. \tag{30}$$

This expression is consistent with the expression for the discount bond prices $P(0,T)$ derived in Chapter 3.

We have also shown in Section 3 that the characteristic function of the density p^δ can be obtained from $\widetilde{V}(t,r;T,\psi)/P(t,T,r)$, which yields

$$\exp\{i\psi M(t,T,r) - \tfrac{1}{2}\psi^2\Sigma(t,T)\}, \tag{31}$$

where

$$\begin{aligned} M(t,T,r) &= r + \int_t^T \theta^*(s) - \sigma^2(T-s)\,ds \\ \Sigma(t,T) &= \sigma^2(T-t). \end{aligned} \tag{32}$$

This characteristic function is the characteristic function of a normal distribution with mean $M(t,T,r)$ and variance $\Sigma(t,T)$. Hence, the fundamental solutions of the partial differential equation (25) imply that under the T-forward-risk-adjusted measure, the spot interest rate r has a normal distribution with mean $M(t,T,r)$ and variance $\Sigma(t,T)$. This result is consistent with the process (24) derived before.

5. Fundamental Solutions for Normal Models

In Chapter 3 we introduced the class of one-factor yield-curve models, where the spot interest rate r is described by the stochastic differential equation

$$dr = \mu(t,r)\,dt + \sigma(t,r)\,dW. \tag{33}$$

An important sub-class of models arises when the spot interest rate is modeled as follows

$$\begin{cases} dy = \big(\theta(t) - a(t)y\big)\,dt + \sigma(t)\,dW \\ \;\; r = F(t,y) \end{cases} \tag{34}$$

The first equation describes the dynamics of an underlying process y. The spot interest rate r is determined from the underlying process via the function $F(t,y)$. The stochastic differential equation defining y is a *linear equation in the narrow sense.*[3] Hence, the process y has a normal

[3] For a discussion of linear stochastic differential equations in the narrow sense, see Arnold (1992), Chapter 8.

distribution. Therefore, we define this class of models as *normal models.* By making the appropriate choice for F, we can show that several well-known models fall in our class of normal models.

For $F(t, y) = y$, we obtain the Hull and White (1990a) model; the choice $F(t, y) = y^2$, leads to the one-factor version of the squared Gaussian model which is discussed in Chapter 6; and for $F(t, y) = e^y$, the Black and Karasinski (1991) model can be obtained. Finally, for the choices $a(t) \equiv 0$, $\sigma(t) \equiv \sigma$ and $F(t, r) = y$, we see that (34) reduces to the Ho-Lee model.

Using the methodology of Vasicek (1977), which is explained in Chapter 3, one can show that, for a normal model, the price $V(t, y; T)$ of an interest rate derivative security at time t which has a payoff at time T satisfies the partial differential equation

$$V_t + \mu^*(t, y)V_y + \tfrac{1}{2}\sigma(t)^2 V_{yy} - F(t, y)V = 0, \tag{35}$$

where

$$\mu^*(t, y) = \big(\theta(t) - a(t)y\big) - \lambda(t, y)\sigma(t) \tag{36}$$

and $\lambda(t, y)$ denotes the market price of risk. If one makes the additional assumption that the market price of risk $\lambda(t, y)$ is of the form[4] $\lambda_1(t) + \lambda_2(t)y$, we can rewrite μ^* as

$$\mu^*(t, y) = \theta^*(t) - a^*(t)y, \tag{37}$$

where

$$\begin{aligned} \theta^*(t) &= \theta(t) - \lambda_1(t)\sigma(t) \\ a^*(t) &= a(t) + \lambda_2(t)\sigma(t). \end{aligned} \tag{38}$$

In this section we will investigate which models in the class of normal models have normally distributed fundamental solutions, since these models are most likely to have a rich analytical structure. We will prove that the only normal models that have normally distributed fundamental solutions, are the models for which $F(t, y)$ is either a linear or a quadratic function in y. This result implies that only models like the Hull and White models, for which F is linear, and models for which F is quadratic have normally distributed fundamental solutions. Both

[4] See also Hull and White (1990a), Footnote 1.

models have a rich analytical structure. All other normal models, for example the Black and Karasinski (1991) model, for which F is an exponential function in y, have fundamental solutions that are *not* normally distributed.

In the remainder of this section we will prove the following proposition:

Proposition. *The only models in the class of normal models that have normally distributed fundamental solutions are models where $F(t, y)$ is either a linear or a quadratic function in y. These models will have a variance term independent of y, and a mean which is linear in y.*

Proof. The prices of interest rate derivatives in a normal model must satisfy the partial differential equation

$$V_t + \big(\theta^*(t) - a^*(t)y\big)V_y + \tfrac{1}{2}\sigma(t)^2 V_{yy} - F(t, y)V = 0. \tag{39}$$

If a model has normally distributed fundamental solutions V^δ, then the Fourier transform $\widetilde{V}$ will be of the form

$$\widetilde{V}(t, y; T, \psi) = P(t, T, y)\exp\{i\psi M(t, T, y) - \tfrac{1}{2}\psi^2\Sigma(t, T, y)\}, \tag{40}$$

where $P(t, T, y)$ is the price of a discount bond, and M and Σ are the mean and variance of the normal distribution respectively. The price of a discount bond P must solve (39) with boundary condition $P(T, T, y) \equiv 1$.

"*Variance independent of y*"
The function $\widetilde{V}$ defined in (40) must solve the partial differential equation (39). The partial derivatives of $\widetilde{V}$ are given by

$$\begin{aligned}
\widetilde{V}_t &= \widetilde{V}\left(\frac{P_t}{P} + i\psi M_t - \tfrac{1}{2}\psi^2\Sigma_t\right)\\
\widetilde{V}_y &= \widetilde{V}\left(\frac{P_y}{P} + i\psi M_y - \tfrac{1}{2}\psi^2\Sigma_y\right)\\
\widetilde{V}_{yy} &= \widetilde{V}\left(\frac{P_{yy}}{P} + i\psi\Big(2\frac{P_y}{P}M_y + M_{yy}\Big)\right.\\
&\qquad\left. -\psi^2\Big(M_y^2 + \frac{P_y}{P}\Sigma_y + \tfrac{1}{2}\Sigma_{yy}\Big) - i\psi^3 M_y\Sigma_y + \tfrac{1}{4}\psi^4\Sigma_y^2\right).
\end{aligned} \tag{41}$$

If we substitute these partial derivatives into (39), we will get only one term containing ψ^4, namely $\frac{1}{2}\sigma(t)^2\widetilde{V}\frac{1}{4}\psi^4\Sigma_y^2$. However, the partial differential equation will only be solved for all ψ, if all terms containing ψ are identically equal to zero. Hence, the term containing ψ^4 can only be zero if $\Sigma_y \equiv 0$, but this implies that the variance term Σ must be independent of y.

"*Mean linear in y*"
For Σ independent of y, the partial derivatives of $\widetilde{V}$ reduce to

$$\begin{aligned}
\widetilde{V}_t &= \widetilde{V}\left(\frac{P_t}{P} + i\psi M_t - \tfrac{1}{2}\psi^2\Sigma_t\right) \\
\widetilde{V}_y &= \widetilde{V}\left(\frac{P_y}{P} + i\psi M_y\right) \qquad (42)\\
\widetilde{V}_{yy} &= \widetilde{V}\left(\frac{P_{yy}}{P} + i\psi\Big(2\frac{P_y}{P}M_y + M_{yy}\Big) - \psi^2 M_y^2\right).
\end{aligned}$$

Substituting these partial derivatives into (39) and simplifying yields

$$\begin{aligned}
-\tfrac{1}{2}\psi^2\big[\Sigma_t + \sigma(t)^2 M_y^2\big] + i\psi\big[M_t + (\theta^*(t) - a^*(t)y)M_y \\
+ \tfrac{1}{2}\sigma(t)^2(2\tfrac{P_y}{P}M_y + M_{yy})\big] = 0. \quad (43)
\end{aligned}$$

Note, that the other partial derivatives of P cancel against F, because the discount bond price is assumed to satisfy the partial differential equation. The term containing ψ^2 will be identically equal to zero only if $\Sigma_t + \sigma(t)^2 M_y^2 \equiv 0$, which is the case if M_y is a function of time only. Therefore M must be of the form

$$M(t,T,y) = M_0(t,T) + M_1(t,T)y, \qquad (44)$$

so we see that the mean M is linear in y.

"*F linear or quadratic in y*"
For this linear form of M, we get that the coefficient of $i\psi$ in (43) will be identically zero only if

$$M_{0t} + M_{1t}y + (\theta^*(t) - a^*(t)y)M_1 + \sigma(t)^2\tfrac{P_y}{P}M_1 \equiv 0. \qquad (45)$$

This is possible only if P_y/P is either independent of y or a linear function in y. Hence, P is either of the form

$$P(t,T,y) = \exp\{A(t,T) + B(t,T)y\} \tag{46}$$

or

$$P(t,T,y) = \exp\{A(t,T) + B(t,T)y + C(t,T)y^2\}. \tag{47}$$

However, a discount bond price P must solve the partial differential equation (39). If $\log P$ is linear in y, then the partial derivative P_t/P is linear in y, and the derivatives P_y/P and P_{yy}/P are independent of y. In this case F can only be linear in y. On the other hand, if $\log P$ is quadratic in y, then the partial derivatives P_t/P and P_{yy}/P are quadratic in y, and the derivative P_y/P is linear in y. In this case F can only be quadratic in y. Hence, the only partial differential equations for which a normally distributed fundamental solution is feasible are the differential equations where F is either a linear or a quadratic function in y, which completes the proof. ■

5

An Analysis of the Hull-White Model

GIVEN the tools we have developed in the previous chapters, we want to analyse some interest rate models which have a rich analytical structure.[1] In Chapter 4 we proved that only normal models where the spot interest rate is a linear or quadratic function of the underlying process y have normally distributed fundamental solutions. Hence, only these models are expected to have a rich analytical structure.

In this chapter we will analyse a model which is linear in the underlying process. It is the model proposed by Hull and White (1994). It is a generalisation of the continuous-time Ho-Lee model we have encountered before. The Ho-Lee model has a very rich analytical structure and is very easy to analyse, however it provides not a very realistic description of the behaviour of interest rates.

The first point on which the Ho-Lee model fails is the fact that it possesses no mean-reversion of interest rates. On the basis of economic theory, there are compelling arguments[2] for the mean-reversion of interest rates. When rates are high, the economy tends to slow down and investments will decline. This implies there is less demand for money and rates will tend to decline. When rates are low, it is relatively cheap to invest, and rates will tend to rise. To obtain a more realistic model, Hull and White added mean-reversion to the Ho-Lee model.

The second weakness of the Ho-Lee model is the fact that the in-

[1] The author would like to thank John Hull for comments and helpful suggestions.

[2] For example, see Hull (1993), Section 15.10.

terest rates are normally distributed, which implies that the interest rates can become negative with positive probability. This problem is not solved in the Hull-White model because interest rates are also normally distributed in the Hull-White model. However, the probability that interest rates become negative is much smaller in the Hull-White than in the Ho-Lee model.

Despite this weakness, the Hull-White model has gained considerable popularity due to its analytical tractability. The Hull-White model can be fitted to the initial term-structure of interest rates analytically. Also prices for discount bonds and options on discount bonds can be valued analytically. The prices of frequently traded instruments like caps, floors, swaptions and options on coupon-bearing bonds can be expressed in terms of options on discount bonds, and can therefore be valued analytically in the Hull-White model.

In this chapter we provide a different derivation of the analytical fomulæ in the Hull-White model. Using the techniques developed in the previous chapters, we find fundamental solutions to the partial differential equation. From the fundamental solutions, we obtain formulæ for the prices of discount bonds. We furthermore obtain an explicit expression for the transition densities under the T-forward-risk-adjusted measure. Using these results we obtain closed form expressions for the prices of options on bonds and various other interest rate derivatives, which are frequently traded in the market.

To calculate prices for other, more complex, interest rate derivatives, the pricing partial differential equation has to be solved numerically. We develop an explicit finite difference method to solve the partial differential equation. This algorithm is different from the algorithm developed by Hull and White (1994).

The rest of the chapter is organised as follows. In Section 1 we introduce the model and find a transformation that simplifies the partial differential equation. In Section 2 we derive analytical formulæ for some interest rate derivatives. In Section 3 we show how the Hull-White model can be fitted to the initial term-structure of interest rates, and how an explicit finite difference algorithm can be set up efficiently to calculate prices for exotic options. We conclude in Section 4 with some examples, where we illustrate the convergence of the explicit finite difference algorithm and compare it to the algorithm of Hull and White (1994).

1. Spot Rate Process

Hull and White (1994) assume that the spot interest rate r follows the process

$$dr = \big(\theta(t) - ar\big)\,dt + \sigma\,dW, \tag{1}$$

where $\theta(t)$ is an arbitrary function of time and a and σ are constants. This model looks very much like the continuous-time Ho-Lee model, except for the additional term $-ar$ in the drift. For $a > 0$ this term adds the desired mean reverting property to the Ho-Lee model.

The Hull-White model can be cast into the framework of Chapter 4 as follows. We can rewrite the spot rate process as

$$\begin{cases} du = \big(\theta(t) - au\big)\,dt + \sigma\,dW \\ \;\; r = u \end{cases} \tag{2}$$

and we see that the Hull-White model is a normal model where the spot interest rate r is a linear function of the underlying process u.

1.1. Partial differential equation

From Chapter 3 and Chapter 4 we know that the value $V(t,r)$ of an interest rate derivative follows the partial differential equation

$$V_t + \mu^*(t,r)V_r + \tfrac{1}{2}\sigma^2 V_{rr} - rV = 0, \tag{3}$$

with

$$\mu^*(t,r) = \theta(t) - ar - \lambda(t,r)\sigma. \tag{4}$$

If one makes the additional assumption that the market price of risk $\lambda(t,r)$ is a function $\lambda(t)$ of time only, we can rewrite the partial differential equation as

$$V_t + \big(\theta^*(t) - ar\big)V_r + \tfrac{1}{2}\sigma^2 V_{rr} - rV = 0, \tag{5}$$

where $\theta^*(t) = \theta(t) - \lambda(t)\sigma$.

1.2. Transformation of variables

Consider the following transformation of variables

$$\begin{aligned} y &= r - \alpha(t) \\ \alpha(t) &= e^{-at}\Big(r_0 + \int_0^t e^{au}\theta^*(u)\,du\Big). \end{aligned} \tag{6}$$

We have chosen $\alpha(t)$ in such a way that $y(0) = 0$.

The price of any interest rate derivative security in terms of the new variable y can be written as $g(t, y)$. We can establish the following relationships between V and g

$$
\begin{aligned}
V(t,r) &\equiv g(t,y) = g\big(t, r - \alpha(t)\big) \\
V_t &= g_t - \big(-a\alpha(t) + \theta^*(t)\big) g_y \\
V_r &= g_y \\
V_{rr} &= g_{yy}.
\end{aligned}
\tag{7}
$$

Substituting these relations into (5) and using $r = y + \alpha(t)$, the partial differential equation reduces to

$$
g_t - ayg_y + \tfrac{1}{2}\sigma^2 g_{yy} - \big(y + \alpha(t)\big) g = 0. \tag{8}
$$

This partial differential equation can be interpreted as the partial differential equation corresponding to an economy where, under the equivalent martingale measure Q^*, the spot interest rate is generated by

$$
\begin{cases}
\;\; dy = -ay\, dt + \sigma\, dW^* \\
r(t) = y(t) + \alpha(t)
\end{cases}
\tag{9}
$$

The advantage of the transformation of variables now becomes apparent. The stochastic process for the underlying variable y is determined only by the volatility parameters a and σ and is independent of the function $\alpha(t)$. The property makes our transformed model much easier to analyse than the original Hull-White (1994) model.

The process that y is assumed to follow is a very simple stochastic process known as an *Ornstein-Uhlenbeck process*[3]. Given a value $y(t)$ at any point in time t, the probability distribution of $y(T)$ for some future time $T > t$ is a normal distribution with mean

$$
e^{-a(T-t)} y(t) \tag{10}
$$

and variance

$$
\frac{\sigma^2}{2a}\left(1 - e^{-2a(T-t)}\right). \tag{11}
$$

---

[3] See Arnold (1992), Chapter 8 for properties of this stochastic process.

The effect of the mean-reversion for $a > 0$ is reflected both in the mean and the variance of the process. For large T the mean of y falls to 0, and the variance tends to $\sigma^2/2a$. This is a much more realistic behaviour than the Ho-Lee model where the variance term $\sigma^2(T-t)$ becomes very large for large T.

For the remainder of this chapter we will work with the transformed model (9). By inverting the transformation of variables (6) the interested reader can, of course, always translate our results in terms of the original variables r and θ^*.

2. Analytical Formulæ

Prices of interest rate derivatives can be calculated in two ways. Using the Feynman-Kac formula we can express solutions to the partial differential equation (8) in terms of the boundary condition $H(T, y(T))$ at time T as

$$g(t, y) = E^*\big(e^{-\int_t^T r(s)\,ds} H(T, y(T)) \,\big|\, \mathcal{F}_t\big), \tag{12}$$

where E^* is the expectation with respect to the process (9). As we showed in Chapter 4, the price of a derivative with a payoff at time T can be more conveniently evaluated using the T-forward-risk-adjusted measure Q_T. The price can then be expressed as

$$g(t, y) = P(t, T, y) E^{Q_T}\big(H(T, y(T)) \,\big|\, \mathcal{F}_t\big), \tag{13}$$

where $P(t, T, y)$ denotes the price of a discount bond with maturity T at time t. A convenient way to determine the discount bond price $P(t, T, y)$ and the distribution of y under the T-forward-risk-adjusted measure is to use the Fourier transform $\tilde{g}$ of the fundamental solutions g^δ.

2.1. Fundamental solutions

To find the Fourier Transform $\tilde{g}$ for the Hull-White model, we have to solve (8) with respect to the boundary condition $\tilde{g}(T, y; T, \psi) = e^{i\psi y}$. From the proof of Chapter 4 we know for a model where the spot interest rate is a linear function of the underlying process the function $\tilde{g}$ must be of the form

$$\tilde{g}(t, y; T, \psi) = \exp\{A(t; T, \psi) + B(t; T, \psi) y\}. \tag{14}$$

Substituting this functional form into (8) and rearranging terms yields

$$y(B_t - aB - 1) + A_t + \tfrac{1}{2}\sigma^2 B^2 - \alpha(t) = 0, \tag{15}$$

which is solved if A and B solve the following system of ordinary differential equations

$$\begin{cases} B_t - aB - 1 = 0 \\ A_t + \frac{1}{2}\sigma^2 B^2 - \alpha(t) = 0 \end{cases} \tag{16}$$

subject to the boundary conditions $A(T;T,\psi) = 0$ and $B(T;T,\psi) = i\psi$. The solution for A and B is given by

$$\begin{aligned} B(t;T,\psi) &= i\psi e^{-a(T-t)} - \frac{1 - e^{-a(T-t)}}{a} \\ A(t;T,\psi) &= \frac{\sigma^2}{2a^3}\left(a(T-t) - 2(1 - e^{-a(T-t)}) + \tfrac{1}{2}(1 - e^{-2a(T-t)})\right) \\ &\quad - i\psi \frac{\sigma^2}{2a^2}(1 - e^{-a(T-t)})^2 - \tfrac{1}{2}\psi^2 \frac{\sigma^2}{2a}(1 - e^{-2a(T-t)}) \\ &\quad - \int_t^T \alpha(s)\, ds. \end{aligned} \tag{17}$$

Substituting these expressions for A and B into (14) yields

$$\tilde{g}(t,y;T,\psi) = \exp\{A(t,T) - B(t,T)y + i\psi M(t,T,y) - \tfrac{1}{2}\psi^2 \Sigma(t,T)\}, \tag{18}$$

with

$$\begin{aligned} A(t,T) &= \frac{\sigma^2}{2a^3}\Big(a(T-t) - 2(1 - e^{-a(T-t)}) \\ &\qquad\qquad + \tfrac{1}{2}(1 - e^{-2a(T-t)})\Big) - \int_t^T \alpha(s)\, ds \\ B(t,T) &= \frac{1 - e^{-a(T-t)}}{a} \\ M(t,T,y) &= y e^{-a(T-t)} - \frac{\sigma^2}{2a^2}\left(1 - e^{-a(T-t)}\right)^2 \\ \Sigma(t,T) &= \frac{\sigma^2}{2a}\left(1 - e^{-2a(T-t)}\right). \end{aligned} \tag{19}$$

In Chapter 4 we demonstrated that the Fourier transform $\tilde{g}$ is the product of the discount bond price and the characteristic function of the probability density function under the T-forward-risk-adjusted measure. Hence, the discount bond price is given by

$$P(t,T,y) = \exp\{A(t,T) - B(t,T)y\}. \tag{20}$$

The remaining terms in $\tilde{g}$ can be recognised as the characteristic function of a normal distribution with mean M and variance Σ. Hence, the probability density function $p^{Q_T}(t,y;T,z)$ for $y(T)$ under the T-forward-risk-adjusted measure is a normal probability density function with mean $M(t,T,y)$ and variance $\Sigma(t,T)$. If we compare this mean and variance to the mean and variance of the process y given in (10) and (11), we see that the variances are the same, but that the mean has changed.

For $a = 0$ the Hull-White model reduces to the continuous-time Ho-Lee model. It is left to the reader to verify that all the results derived above reduce to the formulæ for the Ho-Lee model derived in Chapters 3 and 4 when we take the limit $a \to 0$.

2.2. Option prices

Let us consider the price of a European call option on a discount bond. Let $\mathbf{C}(t,T,\mathcal{T},K,y)$ denote the price at time t of a call option that gives at time T the right to buy a discount bond with maturity $\mathcal{T}$ for a price K, with $t < T < \mathcal{T}$. Suppose that at time T the value of $y(T)$ is equal to z, then the payoff $H(T,z)$ of this option is equal to $\max\{P(T,\mathcal{T},z)-K,0\}$. The payout of the option is non-zero if

$$z < \frac{A(T,\mathcal{T}) - \log K}{B(T,\mathcal{T})}. \tag{21}$$

Under the T-forward-risk-adjusted measure Q_T the price of the call option can be expressed as

$$\mathbf{C}(t,T,\mathcal{T},K,y) = E^{Q_T}\left(\max\{P(T,\mathcal{T},y(T)) - K, 0\} \;\middle|\; \mathcal{F}_t\right). \tag{22}$$

Given the expressions we have derived for the price of the discount bond and the probability distribution p^{Q_T} of $y(T)$ under the T-forward-risk-adjusted measure, this expectation can be written as

$$\int_{-\infty}^{\frac{A-\log K}{B}} \frac{e^{A(T,\mathcal{T})-B(T,\mathcal{T})z} - K}{\sqrt{2\pi\Sigma(t,T)}} \exp\{-\tfrac{1}{2}\tfrac{(z-M(t,T,y))^2}{\Sigma(t,T)}\}\, dz. \tag{23}$$

Some calculation will confirm that this integral can be expressed in terms of cumulative normal distribution functions $N(.)$ as follows:

$$\mathbf{C} = P(t,\mathcal{T},y)N(h) - P(t,T,y)KN(h-\Sigma_P), \tag{24}$$

where

$$\begin{aligned} \Sigma_P &= B(T,\mathcal{T})\sqrt{\Sigma(t,T)} \\ h &= \frac{\log\bigl(P(t,\mathcal{T},y)/P(t,T,y)K\bigr)}{\Sigma_P} + \tfrac{1}{2}\Sigma_P. \end{aligned} \tag{25}$$

The value of a put option $\mathbf{P}$ can be derived in a similar fashion, and can be expressed as

$$\mathbf{P} = P(t,T,y)KN(-h+\Sigma_P) - P(t,\mathcal{T},y)N(-h). \tag{26}$$

2.3. Prices for other instruments

With the analytic formulæ for discount bonds and call- and put-options on discount bonds, prices can be calculated for the following interest rate derivatives that are currently traded:

- **Coupon bearing bonds** can be priced as a portfolio of discount bonds: one discount bond for every coupon payed plus one discount bond for the principal.
- **Swaps** can be split in a fixed and a floating leg. The floating leg is always worth par[4], the fixed leg can be viewed as a portfolio of discount bonds.
- A **European option on a coupon bearing bond** is an option on a portfolio of discount bonds. Jamshidian has shown that this option can be decomposed as a portfolio of options on discount bonds. This is a non-trivial procedure, for references see e.g. Jamshidian (1989), Hull and White (1990a) or Hull (1993), Chapter 15.
- **Caps** and **floors** can also be decomposed as portfolios of puts and calls on discount bonds. This is explained in the appendix.
- Since the floating leg in a swap is always worth par, a **swaption** can be viewed as an option on the fixed leg with a par strike. Since the

[4] Assuming, of course, the next floating rate has not been set yet; if it has been set, the next floating rate payment is known, and can be viewed as a discount bond.

fixed leg is a portfolio of discount bonds, we see that Jamshidian's decomposition into a portfolio of options on discount bonds can be used again.

For valuing American-style instruments, and various other more complex interest rate derivatives no analytic formulas are available. Prices of these instruments can be calculated numerically using a finite difference approach. This will be explained in the next section.

3. Implementation of the Model

In this section we show how the Hull-White model we are considering can be fitted to the initial term-structure of interest rates by choosing $\alpha(t)$, and hence $\theta^*(t)$, such that the initial discount bond prices $P(0,T)$ are priced correctly. To calculate prices for derivatives for which no analytic formulæ are available, we develop an explicit finite difference method for the Hull-White model.

This explicit finite difference algorithm is different from the algorithm of Hull and White (1994). The most important difference is that our explicit finite difference algorithm uses an analytical formula for $\alpha(t)$ that is obtained from fitting the model analytically to the initial term-structure of interest rates. Hull and White fit their tree numerically to the term-structure before they start to calculate prices.

3.1. Fitting the model to the initial term-structure

The initial term-structure of interest rates is given by the prices of the discount bonds $P(0,T)$ at time $t=0$. Given the formula for the discount bond price we have to solve

$$\log P(0,T) = A(0,T). \tag{27}$$

Substituting the definition for A given in (19) and taking derivatives with respect to T and simplifying yields

$$\alpha(T) = -\frac{\partial}{\partial T}\log P(0,T) + \frac{\sigma^2}{2a^2}\left(1-e^{-aT}\right)^2. \tag{28}$$

The same result can be derived in a more elegant way using the T-forward-risk-adjusted measure. In Chapter 4 we derived the result that the T-forward rate is the expected value of the spot interest rate at time T under the T-forward-risk-adjusted measure. Hence, we have

$$f(0,T) = E^{Q_T}\big(r(T)\big) = E^{Q_T}\big(y(T) + \alpha(T)\big). \tag{29}$$

Seen from time $t = 0$ the process $y(T)$ has a normal distribution with mean $M(0,T,0)$ under the measure Q_T and we obtain immediately

$$\alpha(T) = f(0,T) + \frac{\sigma^2}{2a^2}\big(1 - e^{-aT}\big)^2, \tag{30}$$

where we have used the expression for M given in (19).

3.2. Transformation of variables

In Section 1 we derived a transformation of variables that made the process y independent of the initial term-structure of interest rates. In this section we show that with an additional transformation of variables we can make the partial differential equation independent of the function $\alpha(t)$. If we define the function $h(t,y)$ as

$$h(t,y) = e^{\int_t^T \alpha(u)\,du} g(t,y), \tag{31}$$

we get the following partial differential equation for h

$$h_t - ayh_y + \tfrac{1}{2}\sigma^2 h_{yy} - yh = 0 \tag{32}$$

and we see that this partial differential equation has no longer terms dependent on t.

Solving (32) numerically, one can calculate values for $h(t,y)$. The value $V(t,r)$ of an interest rate derivative security is obtained from $h(t,y)$ via

$$V(t,r) = e^{-\int_t^T \alpha(u)\,du} h\big(t, r - \alpha(t)\big). \tag{33}$$

Using the analytic formula for $\alpha(t)$ given in (28), the integral of α can be calculated as

$$\int_t^T \alpha(u)\,du = -\log\frac{P(0,T)}{P(0,t)} + \frac{\sigma^2}{2a^3}\Big(a(T-t) - 2(e^{-at} - e^{-aT}) + \tfrac{1}{2}(e^{-2at} - e^{-2aT})\Big) \tag{34}$$

3.3. Trinomial tree

Suppose we construct a grid with steps Δy along the y-axis, and steps Δt along the t-axis. A node (i,j) on the grid is a point where $t = i\Delta t$ and $y = j\Delta y$. On this grid we can calculate in every node a value h_{ij}, that will be an approximation of the "true" value $h(i\Delta t, j\Delta y)$. In the explicit finite difference approach the partial derivatives of h are approximated as follows[5]

$$\begin{aligned} h_{yy} &\approx \frac{1}{(\Delta y)^2}\{h_{i+1,j+1} - 2h_{i+1,j} + h_{i+1,j-1}\} \\ h_y &\approx \frac{1}{2\Delta y}\{h_{i+1,j+1} - h_{i+1,j-1}\} \\ h_t &\approx \frac{1}{\Delta t}\{h_{i+1,j} - h_{ij}\} \end{aligned} \tag{35}$$

To solve the partial differential equation numerically one imposes the restriction that in every node (i,j) the approximations of the partial derivatives satisfy the partial differential equation (32) exactly, so we get after solving for h_{ij}

$$\begin{aligned} h_{ij} = \frac{\Delta t}{1 + j\Delta y \Delta t}\Big\{ &\big(\frac{\sigma^2}{2(\Delta y)^2} - \tfrac{1}{2}aj\big)h_{i+1,j+1} + \\ &\big(-\frac{\sigma^2}{(\Delta y)^2} + \frac{1}{\Delta t}\big)h_{i+1,j} + \\ &\big(\frac{\sigma^2}{2(\Delta y)^2} + \tfrac{1}{2}aj\big)h_{i+1,j-1}\Big\}. \end{aligned} \tag{36}$$

If we set $\Delta y = \sigma\sqrt{3\Delta t}$, the equation reduces to

$$h_{ij} = \frac{1}{1 + j\Delta y \Delta t}\{p_u h_{i+1,j+1} + p_m h_{i+1,j} + p_d h_{i+1,j-1}\} \tag{37}$$

with

$$\begin{aligned} p_u &= \tfrac{1}{6} - \tfrac{1}{2}aj\Delta t \\ p_m &= \tfrac{2}{3} \\ p_d &= \tfrac{1}{6} + \tfrac{1}{2}aj\Delta t. \end{aligned} \tag{38}$$

---

[5] For an introduction to the numerical solution of partial differential equations, see Hull and White (1990b) or Smith (1985).

For $-\frac{1}{3}\frac{1}{a\Delta t} < j < \frac{1}{3}\frac{1}{a\Delta t}$, the numbers p_u, p_m and p_d are all positive and sum to 1, and can be interpreted as trinomial probabilities. For increasing values of $|j|$, the probability of jumping towards $y = 0$ increases. This reflects the mean reverting behaviour of the process $y(t)$. In order to prevent the trinomial probabilities from going negative, we cannot use a finite difference grid that is arbitrarily large. At some level $j^+ < \frac{2}{3}\frac{1}{a\Delta t}$, we want to express h_{ij^+} in terms of h_{i+1,j^+}, h_{i+1,j^+-1} and h_{i+1,j^+-2}. By doing so we avoid using h_{i+1,j^++1} and the grid will remain bounded at j^+. If we use the following approximations of the partial derivatives

$$\begin{aligned} h_{yy} &\approx \frac{1}{(\Delta y)^2}\{h_{i+1,j^+} - 2h_{i+1,j^+-1} + h_{i+1,j^+-2}\} \\ h_y &\approx \frac{1}{2\Delta y}\{3h_{i+1,j^+} - 4h_{i+1,j^+-1} + h_{i+1,j^+-2}\} \\ h_t &\approx \frac{1}{\Delta t}\{h_{i+1,j^+} - h_{ij^+}\}, \end{aligned} \tag{39}$$

we can express h_{ij^+} as

$$h_{ij^+} = \frac{1}{1 + j^+\Delta y\Delta t}\{p_m^+ h_{i+1,j^+} + p_d^+ h_{i+1,j^+-1} + p_{dd}^+ h_{i+1,j^+-2}\} \tag{40}$$

with

$$\begin{aligned} p_m^+ &= \tfrac{7}{6} - \tfrac{3}{2}aj^+\Delta t \\ p_d^+ &= -\tfrac{1}{3} + 2aj^+\Delta t \\ p_{dd}^+ &= \tfrac{1}{6} - \tfrac{1}{2}aj^+\Delta t. \end{aligned} \tag{41}$$

These probabilities are all positive for $\frac{1}{6}\frac{1}{a\Delta t} < j^+ < \frac{1}{3}\frac{1}{a\Delta t}$. We can also bound the grid from below at a level $j^- > -\frac{1}{3}\frac{1}{a\Delta t}$. At j^- we get

$$h_{ij^-} = \frac{1}{1 + j^-\Delta y\Delta t}\{p_{uu}^- h_{i+1,j^-+2} + p_u^- h_{i+1,j^-+1} + p_m^- h_{i+1,j^-}\} \tag{42}$$

with probabilities

$$\begin{aligned} p_{uu}^- &= \tfrac{1}{6} + \tfrac{1}{2}aj^-\Delta t \\ p_u^- &= -\tfrac{1}{3} - 2aj^-\Delta t \\ p_m^- &= \tfrac{7}{6} + \tfrac{3}{2}aj^-\Delta t, \end{aligned} \tag{43}$$

which are all positive for $-\frac{1}{3}\frac{1}{a\Delta t} < j^- < -\frac{1}{6}\frac{1}{a\Delta t}$.

Instead of first calculating h_{ij} in every node, and then, calculating V_{ij} from h_{ij}, using (33), we can calculate prices V_{ij} more efficiently by rewriting the differencing scheme as

$$g_{ij} = \frac{e^{-\int_{i\Delta t}^{(i+1)\Delta t} \alpha(u)\,du}}{1 + j\Delta y\Delta t}\big\{p_u g_{i+1,j+1} + p_m g_{i+1,j} + p_d g_{i+1,j-1}\big\}, \qquad (44)$$

where we have used (31). The price V_{ij} of an instrument can be obtained from g_{ij} using $V(t,r) = g\big(t, r - \alpha(t)\big)$, where $\alpha(t)$ can be evaluated using (28). For every i, the expression $\int \alpha(u)\,du$ can be calculated analytically using (34).

The finite difference method outlined above can be implemented as follows. For an instrument with maturity T, and a given number of steps N, we can calculate the step-sizes $\Delta t = T/N$ and $\Delta y = \sigma\sqrt{3\Delta t}$. We set $j^{\max} = \lceil \frac{1}{6}\frac{1}{a\Delta t}\rceil$, which is the first integer value of j for which we can bound the grid at $j^{+} = -j^{-} = j^{\max}$ without creating negative probabilities. Then for $i = 0, \ldots, j^{\max}$ we can build a normal trinomial tree with probabilities p_u, p_m and p_d. For $i = j^{\max} + 1, \ldots, N$ we build a trinomial tree that jumps "inward" at $j = j^{\max}$ and $j = -j^{\max}$, where the adjusted probabilities p^{+} and p^{-} have to be used. The value for a derivative can be calculated by filling $h_{N,j}$ with the payoff $H(N\Delta t, j\Delta y)$ and then calculating backward using the backward recursion formula (44). Due to the fact that the tree jumps inward we only need a boundary condition at time T; "upper" and "lower" boundary conditions do not have to be supplied.

4. Performance of the Algorithm

The analysis of the Hull-White model in this chapter is different from the analysis of Hull and White (1990a) and (1994). However, the analytical formulæ we obtain for the prices of discount bonds and options on discount bonds are consistent with the results of Hull and White.

The explicit finite difference (FD) algorithm we developed in the previous section is different from the algorithm of Hull and White (1994) (HW). Due to the fact that Hull and White fit their tree numerically to the term-structure before they start to calculate prices, the HW algorithm is slower than the FD algorithm. This is confirmed in Table I, where we report the calculation times needed for the two algorithms.

Table I:
Calculation time in milli-seconds on 486DX/66MHz computer of 5yr option on 9yr discount bond with $a = 0.10, \sigma = 0.01$

	European		American	
Steps	HW	FD	HW	FD
10	4.0	4.0	11.1	8.2
20	8.4	7.8	23.8	17.0
30	14.7	12.5	39.6	27.9
40	22.3	17.7	57.2	39.8
50	32.2	23.4	78.7	53.2

Table II:
Prices (in bp) of put options on 9yr discount bond with $a = 0.10, \sigma = 0.01$ and a zero-curve given by $Z(T) = 0.08 - 0.05e^{-0.18T}$

		European		American*	
Mat/Strike	Steps	HW	FD	HW	FD
$T = 3$yr	10	199	197	200	198
	20	196	195	197	196
	30	194	193	195	194
$K = 0.63$	40	193	193	194	194
	50	193	193	194	194
	Anal.	193			
$T = 5$yr	10	141	138	149	145
	20	139	138	147	145
	30	138	137	146	145
$K = 0.72$	40	138	137	145	145
	50	137	137	145	144
	Anal.	136			
$T = 7$yr	10	103	101	120	116
	20	100	99	117	115
	30	99	98	115	114
$K = 0.85$	40	98	98	115	114
	50	98	97	114	113
	Anal.	97			

* American style put option gives on early exercise at time t the right to sell a discount bond with maturity $t + 9 - T$ years for a price of K.

We see that the FD algorithm is up to 40% faster than the HW algorithm.

The difference in calculation times between the two algorithms can be largely reduced by enhancing the original Hull and White (1994) algorithm by also fitting it analytically to the initial term-structure of interest rates.

Another difference between the FD algorithm and Hull and White's is the fact that we build a tree for the spot interest rate r, while Hull and White build a tree for the (continuous) Δt-interest rate. This means that all analytical formulæ have to be re-expressed in terms of the Δt rate[6] before they can be used to calculate prices in a Hull-White tree.

In Table II we compare the convergence of both algorithms for European and American-style put options with different maturities on a 9-year discount bond. The initial term-structure of interest rates is given via the zero-curve $Z(T)$. The prices of the discount bonds $P(0,T)$ can be calculated from the zero-curve via $P(0,T) = \exp\{-Z(T)T\}$. We see that both the HW algorithm and the FD algorithm converge very fast. Convergence to within 1 basispoint (bp) is reached with a tree that takes 50 time steps. The FD algorithm converges slightly faster than the HW algorithm, but the difference is not very large.

[6] The spot interest rate can be expressed in terms of the Δt-rate using the identity $P(t, t+\Delta t, r) = \exp\{-r_{\Delta t}\Delta t\}$, where $r_{\Delta t}$ denotes the continuous compounded Δt-interest rate.

Appendix

In this appendix we show how the value of a cap or a floor can be expressed as a string of puts or calls on discount bonds respectively.

As is explained in Hull (1993) Chapter 15, a cap contract with strike X on the 3-months interest rate is a portfolio of options on the quarterly interest payments that have to be made. The individual options are known as *caplets*. At the beginning of a quarter, the (discrete) 3-months interest rate R_3 is observed. Hence, at the end of the quarter the writer of the cap has to make a payment of

$$L\Delta t \max\{R_3 - X, 0\}, \tag{45}$$

where L is the principal and Δt is the length of the quarter in days divided by 360.[7] At the beginning of the quarter, this payoff has a value of

$$L\frac{\Delta t}{1+\Delta t R_3} \max\{R_3 - X, 0\}. \tag{46}$$

A few lines of algebra shows that this is equal to

$$L(1+\Delta t X) \max\left\{\frac{1}{1+\Delta t X} - \frac{1}{1+\Delta t R_3}, 0\right\}. \tag{47}$$

The expression $1/(1+\Delta t R_3)$ is equal to the value of the 3-months discount bond $P(T, \mathcal{T}, r)$, where T and $\mathcal{T}$ denote the start- and end-date, respectively, of the quarter under consideration. Hence, the value of the caplet is equal to $L(1+\Delta t X)$ times the value of a put option on the 3-months discount bond with strike $1/(1+\Delta t X)$. The value of a floor can be derived using similar arguments.

[7] This is the Act/360 daycount convention, which is used for 3-months US-dollar interest rates.

6

An Analytically Tractable Model with Positive Interest Rates

AFTER the analysis in Chapter 5 of the Hull-White model, where the spot interest rate is a linear function of the underlying process, we turn our attention to a model where the spot interest rate is a quadratic function of the underlying process.

In Chapter 5 we showed that the Hull-White (1994) model has a rich analytical structure. Analytical formulæ for the prices of discount bonds and options on discount bonds can be obtained in this model. With these formulæ, prices for frequently traded instruments like caps, floors and options on coupon bonds can be calculated analytically. Also, efficient numerical procedures exist for calculating the prices of derivatives for which no analytical formulæ exist. A major disadvantage of the Hull-White approach, however, lies in the fact that negative interest rates can occur, due to the fact that interest rates are normally distributed.

This problem can be circumvented by assuming that the spot interest rate is a quadratic function of the underlying process. This type of models is known as *squared Gaussian models*.

In this chapter, we show that for the one-factor squared Gaussian model an analytical structure as rich as in the Hull-White (1994) model can be obtained, with the additional advantage that the interest rates never become negative. Squared Gaussian models were first studied by Beaglehole and Tenney (1991) and Jamshidian (1993), these papers show that squared Gaussian models possess some analytical structure.

Using a different approach, we are able to carry the analysis of the one-factor squared Gaussian model considerably further. We provide analytical formulæ for the prices of discount bonds and options on discount bonds, and we show how the model can be fitted analytically to the initial term-structure. In the one-factor squared Gaussian model, the prices of options on discount bonds can be expressed in terms of cumulative normal distribution functions. Also we show that prices of other derivatives can be calculated efficiently in this model using a trinomial tree.

The rest of the chapter is organised as follows. In Section 1 we introduce the model and find a transformation that simplifies the partial differential equation. How we obtain fundamental solutions is explained in Section 2, where we also derive analytical formulæ for some interest rate derivatives. In Section 3 we show how our model can be fitted to the initial term-structure, and how a trinomial tree can be set up efficiently to calculate prices for exotic options. We conclude with some examples, illustrating the convergence of the trinomial tree algorithm.

1. Spot Rate Process

In this chapter we analyse the one-factor squared Gaussian model

$$\begin{cases} du = \big(\theta(t) - au\big)\, dt + \sigma\, dW \\ \ \ r = u^2 \end{cases} \tag{1}$$

where a and σ are constants, and $\theta(t)$ is an arbitrary function of time. This is a normal model where the spot interest rate r is given by the square of the underlying process, which ensures that the spot interest rate never goes negative. For the general n-factor model Jamshidian (1993) explains how prices of discount bonds can be obtained numerically and how the n-factor model can be fitted to the initial term-structure by numerically solving a system of ordinary differential equations. Furthermore, he demonstrates that the prices of options on discount bonds can be expressed in terms of non-central chi-square distribution functions. Jamshidian's analysis for the n-factor model is valid for this model also. However, using a different approach, we will provide analytical formulæ for the prices of discount bonds and we demonstrate that prices of interest rate derivatives can be expressed in terms of cumulative normal distribution functions.

1.1. Partial differential equation

As was shown in Chapters 3 and 4, the price $V(t,u)$ at time t of an interest rate derivative security satisfies the partial differential equation

$$V_t + \big(\theta^*(t) - au\big)V_u + \tfrac{1}{2}\sigma^2 V_{uu} - u^2 V = 0, \tag{2}$$

where $\theta^*(t) = \theta(t) - \lambda(t)\sigma$. To make this partial differential equation easier to solve we employ a transformation of variables similar to the one used in Chapter 5

$$\begin{aligned} y &= u - \alpha(t) \\ \alpha(t) &= e^{-at}\left(\sqrt{r_0} + \int_0^t e^{as}\theta^*(s)\,ds\right). \end{aligned} \tag{3}$$

The price of any interest rate derivative security in terms of the new variable y can be written as $h(t,y) = V\big(t, y + \alpha(t)\big)$. In terms of y all prices follow the partial differential equation

$$h_t - ayh_y + \tfrac{1}{2}\sigma^2 h_{yy} - \big(y + \alpha(t)\big)^2 h = 0. \tag{4}$$

The transformed equation (4) can be interpreted as the partial differential equation belonging to an economy where, under the equivalent martingale measure Q^*, the spot interest rate r is generated via

$$\begin{cases} dy = -ay\,dt + \sigma\,dW^* \\ \;\;r = \big(y + \alpha(t)\big)^2 \end{cases} \tag{5}$$

Given a value $y(t)$ at any point in time t, the probability distribution of $y(T)$ for some future time $T > t$ is a normal distribution with mean

$$e^{-a(T-t)}y(t) \tag{6}$$

and variance

$$\frac{\sigma^2}{2a}\left(1 - e^{-2a(T-t)}\right). \tag{7}$$

From a pricing point of view, the economy (1) with market price of risk $\lambda(t)$ and the "risk-neutral" economy (5) are indistinguishable, because the only observed variable in both economies is the spot interest rate r which has the same stochastic behaviour in both economies. It will make no difference for the price of a derivative whether we view it as a function of u or y. Therefore, we use in the remainder of this chapter the transformed partial differential equation (4), which corresponds to the risk-neutral economy (5).

2. Analytical Formulæ

Prices of interest rate derivatives can be calculated in two ways. Using the Feynman-Kac formula we can express solutions to the partial differential equation (4) in terms of the boundary condition $H(T, y(T))$ at time T as

$$h(t,y) = E^*\big(e^{-\int_t^T r(s)\,ds} H(T, y(T)) \,\big|\, \mathcal{F}_t\big), \tag{8}$$

where E^* is the expectation with respect to the process (5). As we showed in Chapter 4, the price of a derivative with a payoff at time T can be more conveniently evaluated using the T-forward-risk-adjusted measure Q_T. The price can then be expressed as

$$h(t,y) = P(t,T,y) E^{Q_T}\big(H(T, y(T)) \,\big|\, \mathcal{F}_t\big), \tag{9}$$

where $P(t,T,y)$ denotes the price of a discount bond with maturity T at time t. A convenient way to determine the discount bond price $P(t,T,y)$ and the distribution of y under the T-forward-risk-adjusted measure is to use the Fourier transform $\tilde{h}$ of the fundamental solutions h^δ.

2.1. Fundamental solutions

To find the function $\tilde{h}$ for the squared Gaussian model, we have to solve (4) with respect to the boundary condition $\tilde{h}(T, y; T, \psi) = e^{i\psi y}$. From the proof of Chapter 4 we know that, for a model where the spot interest rate is a quadratic function of the underlying process, the function $\tilde{h}$ must be of the form

$$\tilde{h}(t, y; T, \psi) = \exp\{A(t; T, \psi) - B(t; T, \psi) y - C(t; T, \psi) y^2\}. \tag{10}$$

A function $\tilde{h}$ of this form satisfies (4) and the boundary condition $e^{i\psi y}$ if $A(t;T,\psi)$, $B(t;T,\psi)$ and $C(t;T,\psi)$ satisfy the following system of ordinary differential equations

$$\begin{cases} C_t - 2aC - 2\sigma^2 C^2 + 1 = 0 \\ B_t - aB - 2\sigma^2 BC + 2\alpha(t) = 0 \\ A_t + \frac{1}{2}\sigma^2 B^2 - \sigma^2 C - \alpha(t)^2 = 0 \end{cases} \tag{11}$$

with $A(T;T,\psi) = C(T;T,\psi) = 0$ and $B(T;T,\psi) = -i\psi$. The solution to this system, with respect to the boundary conditions, is given by

$$\begin{aligned}
C(t;T,\psi) &= C(t,T) \\
B(t;T,\psi) &= B(t,T) - i\psi D(t,T) \\
A(t;T,\psi) &= A(t,T) + \int_t^T \tfrac{1}{2}\sigma^2\bigl(-2i\psi D(s,T)B(s,T) - \psi^2 D(s,T)^2\bigr)\,ds,
\end{aligned} \tag{12}$$

where

$$\begin{aligned}
D(t,T) &= \frac{2\gamma e^{\gamma(T-t)}}{(a+\gamma)e^{2\gamma(T-t)} + (\gamma - a)} \\
C(t,T) &= \frac{e^{2\gamma(T-t)} - 1}{(a+\gamma)e^{2\gamma(T-t)} + (\gamma - a)} \\
B(t,T) &= 2\int_t^T \frac{e^{\gamma s}\bigl((a+\gamma)e^{2\gamma(T-s)} + (\gamma - a)\bigr)}{e^{\gamma t}\bigl((a+\gamma)e^{2\gamma(T-t)} + (\gamma - a)\bigr)}\alpha(s)\,ds \\
A(t,T) &= \int_t^T \tfrac{1}{2}\sigma^2 B(s,T)^2 - \sigma^2 C(s,T) - \alpha(s)^2\,ds \\
\gamma &= \sqrt{a^2 + 2\sigma^2}.
\end{aligned} \tag{13}$$

Substituting $A(t;T,\psi)$, $B(t;T,\psi)$ and $C(t;T,\psi)$ into (10) yields

$$\begin{aligned}
\tilde{h}(t,y;T,\psi) = \exp\bigl\{&A(t,T) - B(t,T)y - C(t,T)y^2 \\
&+ i\psi M(t,T,y) - \tfrac{1}{2}\psi^2\Sigma(t,T)\bigr\},
\end{aligned} \tag{14}$$

with

$$\begin{aligned}
M(t,T,y) &= D(t,T)y - \int_t^T \sigma^2 D(s,T)B(s,T)\,ds \\
\Sigma(t,T) &= \int_t^T \sigma^2 D(s,T)^2\,ds = \sigma^2 C(t,T).
\end{aligned} \tag{15}$$

For $\psi = 0$ the boundary condition for $\tilde{h}$ reduces to the boundary condition of the discount bond. Hence, the price of a discount bond is given by

$$P(t,T,y) = \tilde{h}(t,y;T,0) = \exp\{A(t,T) - B(t,T)y - C(t,T)y^2\}. \tag{16}$$

The terms containing ψ in (14) are easily recognised as the characteristic function of a normal distribution.[1] The probability density functions p^{Q_T} of y under the T-forward-risk-adjusted measure are, therefore, equal to a normal probability density function

$$p^{Q_T}(t,y;T,z) = \frac{1}{\sqrt{2\pi\Sigma(t,T)}} \exp\left\{ -\frac{1}{2} \frac{\big(z - M(t,T,y)\big)^2}{\Sigma(t,T)} \right\}, \tag{17}$$

with mean $M(t,T,y)$ and variance $\Sigma(t,T)$. Note, that this mean and variance are different from the mean and variance of the process y given in (6) and (7).

Once we have explicitly determined prices for discount bonds, and have obtained the distribution of y under the T-forward-risk-adjusted measure, the price of an interest rate derivative with payoff $H\big(T, y(T)\big)$ at time T can be calculated via (9) as

$$\begin{aligned} h(t,y;T) &= P(t,T,y) E^{Q_T}\left(H(T,y(T)) \;\middle|\; \mathcal{F}_t\right) \\ &= P(t,T,y) \int_{-\infty}^{+\infty} H(T,z) p^{Q_T}(t,y;T,z)\, dz, \end{aligned} \tag{18}$$

where z goes through all the possible values of $y(T)$.

2.2. Option prices

Let us consider the price of a European call option on a discount bond. Let $\mathbf{C}(t,T,\mathcal{T},K,y)$ denote the price at time t of a call option that gives at time T the right to buy a discount bond with maturity $\mathcal{T}$ for a price K, with $t < T < \mathcal{T}$. Suppose that at time T the value of $y(T)$ is equal to z, then the payoff $H(z)$ of this option is equal to $\max\{P(T,\mathcal{T},z) - K, 0\}$. The payout of the option is non-zero if

$$\log P(T,\mathcal{T},z) = A(T,\mathcal{T}) - B(T,\mathcal{T})z - C(T,\mathcal{T})z^2 > \log K. \tag{19}$$

This is true if[2]

$$l = \frac{-B(T,\mathcal{T}) - \sqrt{d}}{2C(T,\mathcal{T})} < z < \frac{-B(T,\mathcal{T}) + \sqrt{d}}{2C(T,\mathcal{T})} = h, \tag{20}$$

[1] See, e.g. Lukacs (1970).

[2] Note that $C(T,\mathcal{T})$ is positive for all $T < \mathcal{T}$.

where

$$d = B(T,\mathcal{T})^2 + 4C(T,\mathcal{T})\big(A(T,\mathcal{T}) - \log K\big), \tag{21}$$

provided that the discriminant d is positive. If $d \leq 0$, the payoff of the call option is never positive, and hence the call option price $\mathbf{C}$ is equal to zero. If $d > 0$, we can calculate the price $\mathbf{C}$ by integrating (18) over the region $l < z < h$. The integral can be expressed in terms of cumulative normal distribution functions $N(.)$ as follows

$$\begin{aligned}\mathbf{C}(t,T,\mathcal{T},K,y) &= P(t,\mathcal{T},y)\left[N\left(\frac{h\tau-\nu}{\sqrt{\tau\Sigma(t,T)}}\right) - N\left(\frac{l\tau-\nu}{\sqrt{\tau\Sigma(t,T)}}\right)\right] \\ &- P(t,T,y)K\left[N\left(\frac{h-M(t,T,y)}{\sqrt{\Sigma(t,T)}}\right) - N\left(\frac{l-M(t,T,y)}{\sqrt{\Sigma(t,T)}}\right)\right]\end{aligned} \tag{22}$$

with

$$\begin{aligned}\nu &= M(t,T,y) - B(T,\mathcal{T})\Sigma(t,T) \\ \tau &= 1 + 2C(T,\mathcal{T})\Sigma(t,T).\end{aligned} \tag{23}$$

The price for a put-option on a discount bond can be derived in a similar fashion.

For valuing American-style instruments, and various other exotic interest rate derivatives no analytic formulæ are available. Prices of these instruments have to be calculated numerically.

3. Implementation Of The Model

As was stated in the first section, yield-curve models take the initial term-structure of interest rates as an input, and price all interest rate derivatives off this curve using no-arbitrage arguments. In this section we show how the model we are considering can be fitted to the initial yield-curve by choosing $\alpha(t)$, and hence $\theta^*(t)$, such that the initial discount bond prices $P(0,T)$ are priced correctly. To calculate prices for derivatives for which no analytic formulæ are available, we provide a special trinomial tree approach for the squared Gaussian model.

3.1. Fitting the model to the initial term-structure

In Chapter 4 we proved the relation

$$f(t,T,y) = E^{Q_T}\big(r(T) \bigm| \mathcal{F}_t\big). \tag{24}$$

Hence, under the T-forward-risk-adjusted measure, the instantaneous forward rate for date T is equal to the expected value of the spot interest rate at time T. The spot interest rate is defined as $r = (y+\alpha)^2$, therefore we can express (24) as

$$f(t,T,y) = \Sigma(t,T) + \big(M(t,T,y) + \alpha(T)\big)^2, \tag{25}$$

because, under the T-forward-risk-adjusted measure, y is normally distributed with mean M and variance Σ.

At time $t = 0$ the prices of the forward rates $f(0,T)$ are known for all T and $\alpha(T)$ has to be chosen such that these initial discount bond prices are priced correctly by the model. The value of $\alpha(T)$ can be related to $f(0,T)$ using (25), and we obtain

$$\begin{aligned} f(0,T) &= \Sigma(0,T) + \big(M(0,T,0) + \alpha(T)\big)^2 \\ &= \Sigma(0,T) + \Big(-\int_0^T \sigma^2 D(s,T) B(s,T)\, ds + \alpha(T)\Big)^2, \end{aligned} \tag{26}$$

where we have substituted the definition for M given in (15). If $f(0,T) \geq \Sigma(0,T)$, we can define

$$F(T) = \sqrt{f(0,T) - \Sigma(0,T)}, \tag{27}$$

and get for $\alpha(T)$

$$\alpha(T) - \int_0^T \sigma^2 D(s,T) B(s,T)\, ds = F(T). \tag{28}$$

The function $B(t,T)$ depends also on α, hence, (28) is an integral equation in α. In Appendix A we show that the solution to (28) can be expressed as

$$\alpha(T) = F(T) + 2\int_0^T e^{-a(T-s)} \Sigma(0,s) F(s)\, ds. \tag{29}$$

It is clear that for $f(0,T,0) < \Sigma(0,T)$, the model cannot be fitted to the initial term-structure.

Once $\alpha(t)$ is determined from the initial term-structure, prices for discount bonds and options on discount bonds can be calculated. All pricing formulæ depend on the functions $A(t,T)$ and $B(t,T)$, which can be determined from (13). However, calculating A and B from (13) would involve a numerical integration for every value of $A(t,T)$ or $B(t,T)$ needed. As was shown by Jamshidian (1993), $A(t,T)$ and $B(t,T)$ can be evaluated more efficient, by expressing them in terms of $P(0,T,0)$ and $B(0,T)$. The values for $P(0,T,0)$ are known at $t=0$ for all T, and the values for $B(0,T)$ have to be calculated only once for different values of T and can then be stored. We show in Appendix B how this procedure can be implemented for our model.

3.2. Trinomial tree

For interest rate derivatives for which no analytical formulæ exist, the partial differential equation (4) has to be solved numerically. In the remainder of this section we show how the partial differential equation can be solved using an explicit finite difference method. Then we show how the explicit finite difference method can be implemented efficiently as a trinomial tree. We conclude with an example to illustrate the convergence of the trinomial tree algorithm.

In Chapter 5 we derived an explicit finite difference algorithm for the Hull-White model. We can use the same methodology to derive an explicit finite difference algorithm for the squared Gaussian model. If we choose a grid with spacings Δt and $\Delta y = \sigma\sqrt{3\Delta t}$ we obtain the following backward recursion formula

$$h_{ij} = \tfrac{1}{1+(j\Delta y+\alpha(i\Delta t))^2\Delta t}\left\{p_u h_{i+1,j+1} + p_m h_{i+1,j} + p_d h_{i+1,j-1}\right\} \qquad (30)$$

with

$$\begin{aligned} p_u &= \tfrac{1}{6} - \tfrac{1}{2}aj\Delta t \\ p_m &= \tfrac{2}{3} \\ p_d &= \tfrac{1}{6} + \tfrac{1}{2}aj\Delta t. \end{aligned} \qquad (31)$$

The only difference between this formula, and the recursion scheme of Chapter 5 is the first term. This term reflects the fact that in the squared Gaussian model the spot interest rate is a quadratic function of the underlying process y. Keeping this difference in mind, one can implement a trinomial tree for the squared Gaussian model in exactly the same way as we did for the Hull-White model in Chapter 5.

Table I:
Prices (in bp) of put options on 9yr discount bond with
$a = 0.10, \sigma = 0.03$ and a zero-curve given by $Z(T) = 0.08 - 0.05e^{-0.18T}$

Mat/Strike	Steps	European	American*
$T = 3$yr, $K = 0.6$	20	164	168
	40	161	165
	60	159	164
	80	160	165
	100	160	165
	Anal.	160	
$T = 5$yr, $K = 0.7$	20	159	179
	40	156	177
	60	154	176
	80	154	175
	100	153	175
	Anal.	153	
$T = 7$yr, $K = 0.85$	20	154	197
	40	151	194
	60	150	193
	80	149	193
	100	149	192
	Anal.	148	

* American style put option gives on early exercise at time t the right to sell a discount bond with maturity $t + 9 - T$ yr for a price K.

To analyse the convergence of the trinomial tree algorithm outlined above, we show in Table I prices of European and American-style put options on a 9 year discount bond, for different number of steps. The initial term-structure of interest rates is given via the zero-curve $Z(T)$. The prices of the discount bonds $P(0, T)$ can be calculated from the zero-curve via $P(0, T) = \exp\{-Z(T)T\}$. It is clear from this table, that the trinomial algorithm converges very fast. Convergence within 1 basispoint (bp) is reached within 100 steps, both for the European and the American-style options.

Appendix A

In this appendix we show how the integral equation (28) can be solved. Using (27) for the right-hand side of (28), and using the definitions for D and B given in (13), we get

$$\alpha(T) - \sigma^2 \int_{s=0}^{T} \frac{2\gamma e^{\gamma(T-s)}}{(a+\gamma)e^{2\gamma(T-s)} + (\gamma - a)} \times \\ \int_{u=s}^{T} 2\frac{e^{\gamma u}\Big((a+\gamma)e^{2\gamma(T-u)} + (\gamma-a)\Big)}{e^{\gamma s}\Big((a+\gamma)e^{2\gamma(T-s)} + (\gamma-a)\Big)} \alpha(u)\, du\, ds = F(T). \quad (32)$$

After interchanging the order of integration, we can simplify this expression to

$$\alpha(T) - \frac{2\sigma^2 e^{\gamma T}}{(a+\gamma)e^{2\gamma T} + (\gamma - a)} \int_0^T (e^{\gamma u} - e^{-\gamma u})\alpha(u)\, du = F(T). \quad (33)$$

Readers familiar with literature on integral equations will recognise (33) as a linear Volterra integral equation of the second kind

$$\alpha(T) - \int_0^T K(T,u)\alpha(u)\, du = F(T), \quad (34)$$

with a separable integration kernel $K(T,u)$ equal to

$$K(T,u) = \frac{2\sigma^2 e^{\gamma T}}{(a+\gamma)e^{2\gamma T} + (\gamma - a)}(e^{\gamma u} - e^{-\gamma u}). \quad (35)$$

It is well known (see Hochstadt (1973) or Griffel (1993)) that a linear Volterra integral equation with a continuous and bounded kernel has a unique solution for every continuous function $F(T)$. In this case, where the kernel is separable, the integral equation (33) can be solved as follows. After differentiating (33) with respect to T, we obtain

$$\alpha_T(T) - 2\sigma^2 \tfrac{\gamma e^{\gamma T}((a+\gamma)e^{2\gamma T} + (\gamma-a)) - e^{\gamma T}(a+\gamma)e^{2\gamma T}2\gamma}{((a+\gamma)e^{2\gamma T} + (\gamma - a))^2} \int_0^T (e^{\gamma u} - e^{-\gamma u})\alpha(u)\, du \\ - \frac{2\sigma^2 e^{\gamma T}}{(a+\gamma)e^{2\gamma T} + (\gamma - a)}(e^{\gamma T} - e^{-\gamma T})\alpha(T) = F_T(T). \quad (36)$$

Rewriting (33), the remaining integral can be expressed as

$$\int_0^T (e^{\gamma u} - e^{-\gamma u})\alpha(u)\,du = \Big(\alpha(T) - F(T)\Big)\frac{(a+\gamma)e^{2\gamma T} + (\gamma - a)}{2\sigma^2 e^{\gamma T}}. \tag{37}$$

Substituting this expression into (36) yields, after some calculation

$$\alpha_T(T) + a\alpha(T) = \beta(T)F(T) + F_T(T), \tag{38}$$

where

$$\beta(T) = \gamma - \frac{2\gamma(\gamma - a)}{(a+\gamma)e^{2\gamma T} + (\gamma - a)}. \tag{39}$$

This ordinary differential equation in α has to be solved subject to the boundary condition $\alpha(0) = F(0)$. The solution can be expressed as

$$\alpha(T) = e^{-aT}\left(F(0) + \int_0^T e^{as}\beta(s)F(s)\,ds + \int_0^T e^{as}F_s(s)\,ds\right). \tag{40}$$

Using partial integration, the second integral can be rewritten as

$$\int_0^T e^{as}F_s(s)\,ds = e^{aT}F(T) - F(0) - \int_0^T ae^{as}F(s)\,ds, \tag{41}$$

and we obtain for (40)

$$\alpha(T) = F(T) + e^{-aT}\int_0^T e^{as}\Big(\beta(s) - a\Big)F(s)\,ds. \tag{42}$$

Substituting for β and simplifying, we end up with

$$\alpha(T) = F(T) + 2\int_0^T e^{-a(T-s)}\sigma^2\frac{e^{2\gamma s} - 1}{(a+\gamma)e^{2\gamma s} + (\gamma - a)}F(s)\,ds, \tag{43}$$

which is equal to expression (29). ∎

Appendix B

In this appendix we show how values for $A(T,\mathcal{T})$ and $B(T,\mathcal{T})$ can be calculated efficiently, using methods outlined in Jamshidian (1993).

If an investor sells a discount bond with maturity $\mathcal{T}$ at an earlier time T, he will receive $P(T,\mathcal{T},y(T))$. Therefore, for $t < T < \mathcal{T}$, one can view the discount bond $P(t,\mathcal{T},y)$ as a security that has a payoff of $P(T,\mathcal{T},y(T))$ at time T. Hence, we can write the price of a discount bond $P(t,\mathcal{T},y)$ as

$$P(t,\mathcal{T},y) = P(t,T,y)\widetilde{E}^T\Big(P(T,\mathcal{T},y(T))\;\Big|\;\mathcal{F}_t\Big). \tag{44}$$

Under the T-forward-risk-adjusted measure, the random variable $y(T)$, seen from time t, is normally distributed with mean $M(t,T,y)$ and variance $\Sigma(t,T)$. Using the formula for the price of a discount bond given in (16), we can evaluate (44) as

$$\exp\Big\{A(t,\mathcal{T})-B(t,\mathcal{T})y-C(t,\mathcal{T})y^2\Big\} = \exp\Big\{A(t,T)-B(t,T)y-C(t,T)y^2\Big\}$$
$$\times\frac{\exp\left\{A(T,\mathcal{T})+\frac{\frac{1}{2}B(T,\mathcal{T})^2\Sigma(t,T)-B(T,\mathcal{T})M(t,T,y)-C(T,\mathcal{T})M(t,T,y)^2}{1+2C(T,\mathcal{T})\Sigma(t,T)}\right\}}{\sqrt{1+2C(T,\mathcal{T})\Sigma(t,T)}}. \tag{45}$$

By substituting the expression for M given in (15), and collecting equal powers of y, we obtain the following three identities

$$C(t,\mathcal{T}) - C(t,T) = \frac{C(T,\mathcal{T})D(t,T)^2}{1+2C(T,\mathcal{T})\Sigma(t,T)} \tag{46}$$

$$B(t,\mathcal{T}) - B(t,T) = \frac{B(T,\mathcal{T})D(t,T) - 2C(T,\mathcal{T})D(t,T)M_1(t,T)}{1+2C(T,\mathcal{T})\Sigma(t,T)} \tag{47}$$

$$A(t,\mathcal{T}) - A(t,T) = A(T,\mathcal{T}) - \tfrac{1}{2}\log\Big(1+2C(T,\mathcal{T})\Sigma(t,T)\Big) + \frac{\frac{1}{2}B(T,\mathcal{T})^2\Sigma(t,T)+B(T,\mathcal{T})M_1(t,T)-C(T,\mathcal{T})M_1(t,T)^2}{1+2C(T,\mathcal{T})\Sigma(t,T)}, \tag{48}$$

with

$$M_1(t,T) = \int_t^T \sigma^2 D(s,T)B(s,T)\,ds. \tag{49}$$

These identities hold for all $t < T < \mathcal{T}$. For the special case $t = 0$, we can make the following observations. For all $T > 0$ we have $A(0,T) = \log P(0,T,0)$, furthermore, from (28) we get that $M_1(0,T) = \alpha(T) - F(T)$. Using these relations combined with (47) and (48), we can express $B(T,\mathcal{T})$ and $A(T,\mathcal{T})$ as follows

$$\begin{aligned} B(T,\mathcal{T}) = & \frac{1+2C(T,\mathcal{T})\Sigma(0,T)}{D(0,T)}\Big(B(0,\mathcal{T})-B(0,T)\Big) \\ & + 2C(T,\mathcal{T})\Big(\alpha(T)-F(T)\Big) \end{aligned} \tag{50}$$

$$\begin{aligned} A(T,\mathcal{T}) = & \log\left(\frac{P(0,\mathcal{T},0)}{P(0,T,0)}\right) + \tfrac{1}{2}\log\Big(1+2C(T,\mathcal{T})\Sigma(0,T)\Big) - \\ & \frac{\frac{1}{2}B(T,\mathcal{T})^2\Sigma(0,T)+B(T,\mathcal{T})(\alpha(T)-F(T))-C(T,\mathcal{T})(\alpha(T)-F(T))^2}{1+2C(T,\mathcal{T})\Sigma(0,T)}. \end{aligned} \tag{51}$$

Finally, we can calculate the values of $B(0,T)$ efficiently the following way. Substituting the definition of the forward rates into (25) we obtain

$$-\frac{\partial \log P(t,T,y)}{\partial T} = \Sigma(t,T) + \Big(M(t,T,y)+\alpha(T)\Big)^2. \tag{52}$$

Substituting the formulæ for $P(t,T,y)$ and $M(t,T,y)$ into this equation and collecting equal powers of y leads, again, to three identities. In this way, we get for B the identity

$$B_T(t,T) = 2D(t,T)\Big(\alpha(T)-M_1(t,T)\Big). \tag{53}$$

At $t=0$ this identity simplifies to $B_T(0,T) = 2D(0,T)F(T)$, from which $B(0,T)$ can be calculated as

$$B(0,T) = 2\int_0^T D(0,s)F(s)\,ds. \tag{54}$$

The values for $B(0,T)$ have to calculated only once for different T, and can then be stored.

7

An Empirical Comparison of One-Factor Models

THIS chapter is the last chapter in the first part of this thesis.[1] We have derived the general theory of valuing derivative securities, and we have shown how this theory can be used for valuing interest rate derivatives. We analysed in Chapters 5 and 6 a linear and a squared normal model which both have a rich analytical structure. However, only little attention has been devoted to the empirical validity of these models. In this chapter we address this problem.

Yield-curve models take the initial term-structure of interest rates (and hence bond- or swap-prices) as an input. To estimate the remaining parameters of the yield-curve models we use prices of actively traded US-dollar interest rate caps and floors, which are quoted on a broker screen. To make a comparison between yield-curve models, we have selected the following one-factor yield-curve models: a Hull-White model, a squared Gaussian model and a log-normal model. These models represent the most important solutions for modelling the yield-curve that have been proposed.

We have chosen such a specification for the models that they all have two unknown parameters determining the term-structure of volatilities. Using the observed cap and floor prices, the parameters can be estimated via non-linear least squares. To decide which of the three models provides the best empirical specification, we test the models against each

[1] The author is grateful to Andy Richardson at Intercapital Brokers for graciously providing the data on cap and floor prices. The author is also indebted to Frank de Jong for many comments and helpful suggestions.

other. However, the different yield-curve models are non-nested non-linear models. This implies that the standard t-tests cannot be used as specification tests. Hence, we use the P-test proposed by Davidson and MacKinnon (1993) to discriminate between the different models.

The results of the tests show that the log-normal model is the model that describes the observed cap and floor prices best. However, there is some evidence that a model with a distribution which is even more skewed to the right might provide a better fit.

The rest of this chapter is organised as follows. Section 1 describes the yield-curve models we want to consider. Section 2 describes the econometric approach to estimate and test the models. Section 3 describes the data. Section 4 presents and discusses the empirical results. Section 5 summarises and concludes.

1. Yield-curve Models

The first model we want to consider is the Hull-White (1994) (HW) model, which was analysed in Chapter 5. The spot interest rate r is described in this model by the stochastic differential equation

$$dr = \big(\theta(t) - ar\big)\, dt + \sigma\, dW. \tag{1}$$

The HW model has a rich analytical structure. Analytical expressions for prices of discount bonds can be obtained for this model. Prices for options on bonds can be expressed in terms of cumulative normal distribution functions.

A major disadvantage of the HW model, however, lies in the fact that negative interest rates can occur, due to the assumption that interest rates are normally distributed.

The second model in this study, is a model that precludes the possibility of negative interest rates, without losing the analytic tractability of the HW model. It is the one-factor squared Gaussian (SG) model, studied in Chapter 6

$$\begin{cases} du = \big(\theta(t) - au\big)\, dt + \sigma\, dW \\ \;\; r = u^2 \end{cases} \tag{2}$$

In this model the spot interest rate never becomes negative and has a non-central chi-square distribution with one degree of freedom.

In the SG model analytical expressions for prices of discount bonds can be obtained, also prices of options on discount bonds can be expressed in terms of cumulative normal distribution functions.

Another way to circumvent negative interest rates is to extend the square root model of Cox, Ingersoll and Ross (1985). Several extensions have been proposed, see e.g. Cox, Ingersoll and Ross (1985) and Hull and White (1990a), which allow the square root model to be fitted to the initial term-structure of interest rates but destroy the analytic properties of the square root model. Jamshidian (1995) proposes an extension of the square root model that preserves the analytical properties of the square root model, which he calls the simple square root model. Rogers (1993) shows that simple square root models are equivalent to multi-factor squared Gaussian models. Since we consider only one-factor models in this thesis, we have not included a simple square root model in this study.[2]

The third model we consider is a log-normal model (LN). Log-normal yield-curve models have been developed by Black, Derman and Toy (1990) and later generalised by Black and Karasinski (1991). The version of the log-normal model we consider in this chapter is a restricted version of the Black-Karasinski model due to Hull and White (1994) where the spot interest rate r follows the process

$$d\log r = \big(\theta(t) - a\log r\big)\,dt + \sigma\,dW. \tag{3}$$

Here, the logarithm of the spot interest rate follows a Hull-White process, and is normally distributed. Hence, the spot interest rate itself has a log-normal distribution and the occurrence of negative interest rates is also impossible in the LN model. The largest disadvantage of the LN model is that no analytical formulæ are known for the prices of discount bonds and options on discount bonds. To calculate these prices we have to resort to numerical methods.

---

[2] We have investigated the “exponential decay” model proposed in the paper by Jamshidian (1995), which is an example of a simple square root model. This model cannot be fitted to all term-structures in our sample, also the parameters are very hard to estimate due to the fact that the objective function is very flat around the minimum. For the days the model can be fitted, it fits the observed cap and floor prices about as well as the squared Gaussian model.

Consider the general normal model

$$\begin{cases} dy = -ay\,dt + \sigma\,dW \\ r = F\big(\alpha(t) + y\big) \end{cases} \tag{4}$$

In this general model, the underlying process y follows an Ornstein-Uhlenbeck process. The process y depends only on the volatility parameters a and σ, and is independent of the initial term-structure. A little algebra will reveal that for[3]

$$\alpha(t) = e^{-at}\left(F^{-1}(r_0) + \int_0^t e^{as}\theta(s)\,ds\right) \tag{5}$$

with $F(x) \equiv x$ we obtain the HW model, with $F(x) \equiv x^2$ we obtain the SG model and with $F(x) \equiv e^x$ we obtain the LN model. Since there is a one-to-one relationship between $\alpha(t)$ and $\theta(t)$, it is also possible to use $\alpha(t)$ directly to fit the models to the initial term-structure of interest rates, and obtain $\theta(t)$ via

$$\theta(t) = \frac{d}{dt}\alpha(t) + a\alpha(t), \tag{6}$$

which can be obtained from differentiating (5) with respect to t.

To calculate the prices of caps and floors, we can use for the HW and the SG models the analytic formulæ derived in Chapter 5 and 6, respectively. For the LN model, the prices for caps and floors have to be calculated numerically. In this study we have used the trinomial tree approach outlined in Hull and White (1994).

2. Econometric Approach

The main focus of empirical work in the literature so far has been directed towards models with an endogenous term-structure of interest rates. Using the prices for actively traded (discount) bonds, several authors have estimated the parameters of several models, see e.g. Brown and Dybvig (1986), Longstaff and Schwartz (1992), Chan, Karoly, Longstaff and Sanders (1992) and De Munnik and Schotman (1994).

[3] This is the general form of the transformation of variables used in Chapters 5 and 6.

One of the most interesting papers in this area is Chan, Karoly, Longstaff and Sanders (1992). Using Treasury bill yield data they use the Generalised Method of Moments to estimate and test the model

$$dr = (\alpha - \beta r)\, dt + \sigma r^{\gamma}\, dW. \tag{7}$$

This general specification encompasses several of the endogenous term-structure models. For $\gamma = 0$ the Vasicek, and for $\gamma = \frac{1}{2}$ the Cox-Ingersoll-Ross model is obtained. For $\gamma = 1$ a kind of log-normal model is obtained. Although the stochastic process (7) is very general, it remains unclear whether it is well defined for $\gamma > 1$ (see Rogers (1993)). Their results have to be interpreted with some care. As explained earlier, models with an endogenous term-structure of interest rates cannot be fitted to the initial term-structure, which makes these models less attractive for pricing and managing interest rate derivatives.

For this reason we consider in this chapter yield-curve models, and we adopt a different approach. As was explained in the previous section, the function $\alpha(t)$ (and hence $\theta(t)$) can be determined from the initial term-structure. To determine the parameters a and σ for each model, we fit the models to observed cap and floor prices. Hull and White (1994) suggest that the parameters for their model can be estimated by minimizing the sum of squared differences between the observed cap/floor prices and the prices calculated by the model. However, in the data the observed prices on any given day can range from 1 basispoint to 999 basispoints, which is almost three orders of magnitude. Therefore we have chosen to fit the models to the observed prices in logarithmic terms, which is equivalent to minimizing the relative pricing errors of the models. A fit in logarithmic terms has the additional advantage that the heteroscedasticity of the errors is reduced. Consequently, the econometric model can be written as

$$\boldsymbol{y} = \boldsymbol{g}(\boldsymbol{X}; a, \sigma) + \boldsymbol{\epsilon}, \tag{8}$$

where $\boldsymbol{y}$ is the vector of the logarithms of observed cap and floor prices, and $\boldsymbol{X}$ is a matrix of explanatory variables, which in our case consists of the maturity, strike level and a binary variable indicating cap or floor for every contract.[4] The (vector valued) function $\boldsymbol{g}$ denotes the logarithm

[4] The first three columns of Table II (page 103) contain an example of $\boldsymbol{X}$.

of the price calculated in a model. Estimates $\hat{a}$ and $\hat{\sigma}$ for the parameters are obtained via non-linear least squares. If the errors are independently, identically distributed with zero mean and the function $\boldsymbol{g}$ is sufficiently differentiable, then the non-linear least squares estimates are consistent and have an asymptotical normal distribution. (See e.g. Judge et al. (1982))

The intention of this study is to identify the model that fits the observed cap/floor prices best. Like the option models used for equity and foreign-exchange derivatives, yield-curve models are used as relative valuation tools by banks and institutions. Taking the market prices of actively traded derivatives as given, yield-curve models are calibrated to these prices and are then used to calculate prices and hedge ratios for more complicated (and less liquid) derivatives. For this reason it is important to find a parsimonious model that provides a good fit to observed market prices. Therefore, we do not impose the restriction that the parameters estimates $\hat{a}$ and $\hat{\sigma}$ of any of the models are constant over time. Consequently, we estimated and tested the models on a daily basis using the cap/floor prices observed on any given day. This way of treating the models is consistent with the way yield-curve models are used in practice to price and manage interest rate derivatives.

To determine which of the three yield-curve models describes the data best, we have to choose from the following three hypotheses

$$H_{\mathrm{HW}}: \quad \boldsymbol{y} = \boldsymbol{g}_{\mathrm{HW}}(\boldsymbol{X}; a_{\mathrm{HW}}, \sigma_{\mathrm{HW}}) + \boldsymbol{\epsilon}_{\mathrm{HW}} \tag{9}$$

$$H_{\mathrm{SG}}: \quad \boldsymbol{y} = \boldsymbol{g}_{\mathrm{SG}}(\boldsymbol{X}; a_{\mathrm{SG}}, \sigma_{\mathrm{SG}}) + \boldsymbol{\epsilon}_{\mathrm{SG}} \tag{10}$$

$$H_{\mathrm{LN}}: \quad \boldsymbol{y} = \boldsymbol{g}_{\mathrm{LN}}(\boldsymbol{X}; a_{\mathrm{LN}}, \sigma_{\mathrm{LN}}) + \boldsymbol{\epsilon}_{\mathrm{LN}} \tag{11}$$

Two approaches can be adopted to determine the "best" model. One approach is based on goodness-of-fit criteria like R^2 or the standard error of the regression. These criteria are often easy to compute and intuitively appealing. However, goodness-of-fit criteria have several disadvantages. Using these criteria it is difficult to determine whether one model is "significantly" better than another model, or (even worse) whether all models describe the data badly. In other words, goodness-of-fit criteria do not take into consideration the losses associated with choosing an incorrect model (Judge et al. (1982), Chapter 22.4).

These problems can be avoided by formally testing models against each other. The three models under consideration lead to three hypotheses which are non-linear and non-nested. The testing of non-nested non-linear hypotheses is a hard problem, and no consensus exists which test procedure is to be preferred. For an overview of some test procedures that have been suggested see, for example, Fisher and McAleer (1981), Mizon and Richard (1986). A class of tests which is intuitively appealing and easy to implement, are tests which are based on the principle of artificial nesting. Suppose we have two non-nested hypotheses,

$$H_1 : \quad \boldsymbol{y} = \boldsymbol{g}(\boldsymbol{X};\gamma) + \boldsymbol{\epsilon}_0 \tag{12}$$

$$H_2 : \quad \boldsymbol{y} = \boldsymbol{h}(\boldsymbol{X};\boldsymbol{\delta}) + \boldsymbol{\epsilon}_1 \tag{13}$$

where $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are the parameter-vectors, then we can create an artificial compound model

$$H_{\mathrm{C}} : \quad \boldsymbol{y} = (1-\beta)\boldsymbol{g}(\boldsymbol{X};\gamma) + \beta\boldsymbol{h}(\boldsymbol{X};\boldsymbol{\delta}) + \boldsymbol{\epsilon}. \tag{14}$$

The artificial parameter β has been introduced to nest H_1 and H_2 into H_{C}. The test is then based on the hypothesis $\beta = 0$ or $\beta = 1$. An additional advantage of the artificial compound model is that it can be used as a specification test. If β is significantly different from 0 and 1, it can provide an indication in which direction we have to search for a better model.

The artificial model H_{C} can, in general, not be estimated. If the nesting parameter β approaches either 1 or 0, the parameters $\boldsymbol{\gamma}$ or $\boldsymbol{\delta}$ will not be identified. A solution to this problem (see e.g. Davidson and MacKinnon (1993)) is to replace the parameter vector of the alternative model by a consistent estimate of the parameter vector under the alternative hypothesis. If we take H_2 to be the alternative hypothesis, then we can replace $\boldsymbol{\delta}$ by its non-linear least squares estimate $\hat{\boldsymbol{\delta}}$, and we obtain the alternative model

$$\boldsymbol{y} = (1-\beta)\boldsymbol{g}(\boldsymbol{X};\gamma) + \beta\hat{\boldsymbol{h}} + \boldsymbol{\epsilon}, \tag{15}$$

where $\hat{\boldsymbol{h}} = \boldsymbol{h}(\boldsymbol{X};\hat{\boldsymbol{\delta}})$. If H_1 and H_2 are really non-nested, then both β and $\boldsymbol{\gamma}$ are asymptotically identifiable and have an asymptotically normal distribution. The hypothesis H_1 can be tested, by testing the null hypothesis $\beta = 0$. To implement the test, one can use the t-statistic from the non-linear regression (15). This is known as the J-test.

If we use the Taylor approximation $\boldsymbol{g} \approx \hat{\boldsymbol{g}} + \hat{\boldsymbol{G}}(\boldsymbol{\gamma} - \hat{\boldsymbol{\gamma}})$, where $\hat{\boldsymbol{G}}$ is the matrix of partial derivatives of $\boldsymbol{g}$ with respect to $\boldsymbol{\gamma}$ evaluated at $\hat{\boldsymbol{\gamma}}$, the non-linear regression (15) can be linearized to obtain the Gauss-Newton regression

$$\boldsymbol{y} - \hat{\boldsymbol{g}} = \hat{\boldsymbol{G}}\boldsymbol{c} + b(\hat{\boldsymbol{h}} - \hat{\boldsymbol{g}}) + \boldsymbol{\eta}. \tag{16}$$

An alternative procedure to test H_1 is to use the t-statistic for $b = 0$ from the Gauss-Newton regression. This procedure is much simpler to implement and is asymptotically equivalent to the J-test under H_1 (see Davidson and MacKinnon (1993)) and is called the P-test. It is this test procedure we have implemented.

Since it is just as valid to test H_1 as H_2 against the composite hypothesis H_{C}, we can have four possible outcomes:

i accept H_1 and reject H_2;
ii reject H_1 and accept H_2;
iii accept both hypotheses;
iv reject both hypotheses.

In case *iii* the test indicates that both models are satisfactory, in case *iv* both models are rejected against the artificial compound model and we have to search for another model with more explanatory power. In this case, the value of the nesting parameter can give an indication in which direction we have to look for a better model.

3. Data

We have used data from two sources. Intercapital Brokers provided us with data on cap and floor prices. For 1993 and 1994 they provided us with copies from the Reuters page ICAV at 5:30pm (London time) each trading day. This shows Intercapital's bid and offer quotes for actively traded US-dollar 3 months interest rate caps and floors, totalling 508 trading days. Note that these prices are not closing prices, because US-dollar interest caps and floors are traded continuously on an around the world basis. Each day Intercapital shows quotes for caps and floors with three different strike levels (depending on the level of interest rates) and six different maturities (1,2,3,4,5 and 10 years), giving a total of 36 caps/floors. From the bid and offer quotes mid-market prices were calculated to which the models were fitted.

Figure 1: Standard Error of Regressions

Figure 2: Parameter Estimates

Table I: Summary statistics for parameter estimates
Summary statistics computed over the complete sample of 508 days.

		Mean	St.Dev.	Min.	Q1	Median	Q3	Max.
HW	$\hat{a}$	−0.031	0.076	−0.179	−0.072	−0.043	−0.023	0.426
	$\hat{\sigma}$	0.011	0.002	0.007	0.010	0.011	0.012	0.025
	σ_{LS}	0.206	0.058	0.093	0.161	0.204	0.237	0.531
SG	$\hat{a}$	0.045	0.099	−0.088	−0.016	0.024	0.056	0.463
	$\hat{\sigma}$	0.026	0.005	0.018	0.023	0.026	0.028	0.047
	σ_{LS}	0.142	0.059	0.050	0.133	0.133	0.159	0.535
LN	$\hat{a}$	0.149	0.168	0.000	0.051	0.103	0.160	0.888
	$\hat{\sigma}$	0.254	0.065	0.140	0.212	0.242	0.268	0.481
	σ_{LS}	0.108	0.068	0.023	0.067	0.091	0.122	0.548

Q1 and Q3 denote first and third quartile.

The interest rates for every trading day were obtained from Datastream. For every trading day in the sample we downloaded the overnight rate, the 1,3,6 and 12-months US-dollar money-market rates, and the 2,3,4,5,7 and 10-year US-dollar swap-rates. From these observed rates, 11 continuously compounded zero-rates can be calculated. The complete zero-curve was obtained by log-linear interpolation of the discount bond prices.

4. Empirical Results

As was explained above, the three models under consideration were estimated and tested on a daily basis. In this section we summarise the results for these daily regressions. First we present the parameter estimates and the standard errors of the regressions of the individual models, then we show the results of the pairwise P-tests we conducted.

Table I, Figure 1 and Figure 2 show the results of the 508 daily regressions. Table I contains summary statistics for the parameter estimates of a and σ and the standard error of the regression σ_{LS} for each of the three models. For the complete sample of 508 observations we report the mean and standard deviation of the estimated parameters. Furthermore we show the minimum, first quartile, median, third quartile and maximum value of the estimated parameters over the sample.

Figure 1 shows σ_{LS} for all three models in graphical form. Because all three models have two parameters, this measure is equivalent to R^2.

However, all three models have in general R^2s in excess of 0.99, hence the standard error of the regression σ_{LS} shows the differences between the models more clearly. Over the complete sample, the HW model fits the cap/floor prices worst with an average value for the standard error equal to 0.206. The SG model fits the data better with an average standard error of 0.142. The LN model fits the observed prices best, although towards the end of 1994 the difference between the models becomes less pronounced. The average standard error for the LN model is equal to 0.108. If the standard error of the regression (or R^2) is used as a model selection criterion, we would clearly choose the LN model. However, goodness-of-fit criteria are generally considered to be a poor model selection criterion.

The individual parameter estimates for the three models are shown in graphical form in Figure 2. For every parameter we plotted the estimates and $+2$ and -2 times the asymptotic standard error of the estimates, which gives an indication whether the estimate is significantly different from zero. For all three models, the parameter σ is significantly different from zero, often more than 10 standard deviations away from zero. For the mean-reversion parameter a the picture is entirely different. On the basis of economic theory, there are "compelling" (Hull (1993), Chapter 15.10) arguments for the mean-reversion of interest rates. This should be reflected by the fact that the parameter a is significantly positive. As can be seen from the graphs, a_{HW} is significantly negative for prolonged periods in our sample. The SG model has a mean-reversion parameter which, except for the first few months, is not significantly different from zero. Only the LN model has a mean-reversion parameter that is significantly positive for the larger part of the sample.[5]

The fact that a_{HW} and a_{SG} do not have the correct sign, can be seen as an indication that the HW and the SG model do not describe the empirical steady-state distribution of interest rates adequately. If the empirical steady-state distribution of interest rates has a relatively fat tail on the right, the only way the HW and SG models can mimick this is by increasing the long-term variance, which can be achieved by

[5] Although the LN model is well-defined for $a_{\text{LN}} < 0$, we cannot build a trinomial tree in that case. Therefore, we estimated the LN model with the restriction $a_{\text{LN}} > 0$. However, the restriction was binding for only 21 days in the sample.

Table II: Observed and fitted cap/floor prices on 04-jan-94

			Prices (in bp)			
Mat.	C/F	Strike(%)	Bid–Ask	HW	SG	LN
1yr	C	3.25	53–55	46	46	45
	C	3.50	37–39	32	31	31
	C	3.75	24–27	21	20	20
	F	3.75	10–12	16	15	15
	F	3.50	03–06	8	7	7
	F	3.25	02–04	4	3	2
2yr	C	5.00	45–49	42	44	47
	C	5.50	26–30	21	24	28
	C	6.00	15–18	10	12	16
	F	4.50	72–76	73	74	76
	F	4.00	30–34	34	34	34
	F	3.50	07–10	11	10	9
3yr	C	5.00	138–146	129	134	139
	C	5.50	96–103	82	88	96
	C	6.00	66–73	48	56	66
	F	4.50	90–97	89	91	93
	F	4.00	38–44	43	42	42
	F	3.50	10–13	16	14	12
4yr	C	5.00	270–284	255	262	268
	C	5.50	203–217	178	188	199
	C	6.00	152–165	120	133	147
	F	4.50	104–116	105	107	109
	F	4.00	46–54	52	51	50
	F	3.50	13–19	21	18	15
5yr	C	6.50	196–214	154	174	195
	C	7.00	152–168	107	127	150
	C	7.50	120–134	72	92	115
	F	5.50	336–356	309	321	333
	F	5.00	216–236	204	211	216
	F	4.50	124–138	122	123	124
10yr	C	6.50	636–686	712	707	697
	C	7.00	524–574	580	584	582
	C	7.50	432–482	471	482	487
	F	5.50	480–540	575	548	520
	F	5.00	310–360	419	381	345
	F	4.50	176–216	294	247	207
			$\hat{a}$:	−0.118	−0.060	0.035
			$\hat{\sigma}$:	0.008	0.020	0.206

Table III: Results of pairwise P-tests
Column entries are: # of days; % of sample (508 days)

sign. lvl	acc. HW/rej. SG		rej. HW/acc. SG		acc. HW&SG		reject HW&SG	
10%	0	0%	55	11%	27	5%	426	84%
5%	3	1%	43	8%	14	3%	448	88%
1%	4	1%	40	8%	9	2%	455	90%
	acc. LN/rej. HW		rej. LN/acc. HW		acc. LN&HW		reject LN&HW	
10%	393	77%	10	2%	20	4%	85	17%
5%	335	66%	9	2%	8	2%	156	31%
1%	297	58%	10	2%	4	1%	197	39%
	acc. LN/rej. SG		rej. LN/acc. SG		acc. LN&SG		reject LN&SG	
10%	342	67%	16	3%	73	14%	77	15%
5%	305	60%	20	4%	45	9%	138	27%
1%	273	54%	26	5%	29	6%	180	35%

setting the mean-reversion parameter to a small or negative value.

To get an indication of the fit of the three models, we have reported in Table II the observed and fitted cap and floor prices on January 4, 1994. This day is observation day number 254, and is in the middle of the sample. The first thing to notice is that for longer maturities, the differences between the models becomes more pronounced. For short maturities there is relatively little uncertainty in the interest rates. This means that for short maturities the distribution of interest rates is relatively spiked and can be well approximated by a normal distribution. For longer maturities the uncertainty over future interest rates becomes larger, and the differences between the normal, chi-square and log-normal distribution is prominently reflected in the prices. For longer maturities it is clear that the fit of the LN model is much better than the fit of the HW and the SG model. The differences in skewness of the distributions implied by the three models are reflected by the fact that higher values are assigned to out-of-the-money caps and lower values to out-of-the-money floors as the distributions become more skewed to the right.

To cast the model selection problem into a more formal setting, we have implemented three sets of pairwise P-tests among the three models. The results of these tests are reported in Figure 3 and Table III. Figure 3 shows the estimated nesting parameter b from the P-test regressions, and $+2$ and -2 times the asymptotic standard error of the estimate. As

Figure 3: Pairwise P-test regressions

explained in Section 2, we can consider each of the two models in a P-test as the "null-hypothesis" which can lead to four outcomes of the P-test. Table III summarises how many times each of the outcomes occur at different significance levels for the three pairwise tests conducted.

Both the HW and the SG model are overwhelmingly rejected against the data. The estimations of the nesting parameters indicate consistently that a superior model is obtained by taking two times the SG model minus one time the HW model. This may seem a nonsensical model at first sight. However, by making this combination of models the artificial compound model tries to create a model which is more skewed to the right than the SG model. The results of this P-test point clearly in the direction of a log-normal model. If we test the LN model against the other two models, this is confirmed. In 60 to 70% of the sample the LN model is accepted as the correct model, however for some 30% of the sample all models are rejected. These results indicate that the LN model does a fairly good job in describing the observed cap and floor prices and is certainly the best model of the three models. From the estimated values of the nesting parameters we see that (especially for the first half of the sample) there is still some evidence that a distribution more skewed to the right than the log-normal distribution would provide a still better fit. However, towards the end of the sample, the dataset becomes somewhat less informative and the tests cannot clearly distinguish between the models.

It is interesting to note that these results are consistent with the findings of Chan, Karoly, Longstaff and Sanders (1992), although they use a very different approach. The value for γ Chan et al. find after estimating (7) is approximately 1.5 with a standard error of 0.25, indicating that a log-normal model cannot be rejected but also indicating that a more skewed distribution would be appropriate. However, it remains unclear whether the stochastic process (7) for r is well defined for $\gamma > 1$, see Rogers (1993).

5. Conclusions

In this chapter we have compared three one-factor yield-curve models. We have considered a normal, a squared Gaussian and a log-normal model, which represent the three main types of yield-curve models that have been proposed in the literature. To compare the models we have tested how well the models fit observed US-dollar interest rate caps and floors over a two year sample period. We find that, of the models considered, the log-normal model provides the best fit to the observed cap and floor prices. However, log-normal models have very little analytical properties and all prices must be calculated using numerical methods. Further research into this area might prove fruitful.

Our tests also indicate that a model which implies a distribution of the interest rates which is even more skewed to the right than a log-normal distribution might provide an even better fit to the data. However, it remains unclear whether statistically well defined processes can be designed that possess this property. Much work remains to be done in this area.

Part 2:

Efficient Methods for Valuing and Managing Other Derivative Securities

8

Efficient Calculation of Hedge Parameters in Binomial Models

IT is well known that the Cox, Ross, Rubinstein (1979) model is a very powerful tool for pricing options.[1] Although analytical formulæ are available for certain kinds of options (e.g. the Black-Scholes formula for pricing European options) the binomial model has proven to be very efficient in cases where analytical methods fail (e.g. for American-style put options).

Professional option traders are not only interested in prices, but also in the hedge parameters such as delta, gamma, theta, rho and vega, which are derivatives of the option price function. These "Greeks" are used to evaluate and manage the risks of their option books. Usually the derivatives are approximated by numerical differentiation (finite differences). It is widely known that the three most important derivatives (delta, gamma and theta) can be approximated efficiently by extending the binomial tree. Many believe that this is simply a fast but crude way of calculating these derivatives. The following will show, however, that this is the only method that leads to good approximations and that numerical differentiation (especially where gamma is concerned) might result in approximations that make no sense at all.

We illustrate our point for a European call option where the underlying stock does not pay dividends during the life of the option. This enables us to compare the approximations of the derivatives with the

[1] This chapter is based on the paper "The Binomial Model and the Greeks", by A. Pelsser and T. Vorst, which was published in the Journal of Derivatives (1994).

$$
\begin{tikzcd}[column sep=small, row sep=small]
 & & S_{2,2} \\
 & S_{1,1} \arrow[ur] \arrow[dr] & \\
S \arrow[ur] \arrow[dr] & & S_{2,1} \\
 & S_{1,0} \arrow[ur] \arrow[dr] & \\
 & & S_{2,0}
\end{tikzcd}
$$

Figure 1. Binomial Tree

true values of the Black-Scholes formula.

1. Two Methods to Calculate "Delta" and "Gamma"

Figure 1 gives the well-known binomial tree where $S_{n,j}$ is the stock price in period n and state j. Usually $S_{n,j} = Su^j d^{n-j}$, where u and d are chosen such that the tree gives a good approximation of the underlying stock price process. If $C_{n,j}$ denotes the price of the call option in period n and state$j$, then the option price can be calculated using the backward recursion

$$C_{n,j} = \frac{\pi C_{n+1,j+1} + (1-\pi)C_{n+1,j}}{r}, \tag{1}$$

where r represents the riskless interest rate and $\pi = (r-d)/(u-d)$. The recursion leads to the actual option price C_{00}. We denote this price by $C(S)$, since our tree starts at the stock price S. In a numerical differentiation method, one typically chooses small positive numbers h and k, and constructs new trees with $S_{0,0} = S+h$ and $S_{0,0} = S-k$, respectively. The hedge ratios Δ and Γ are then approximated by

$$
\begin{aligned}
\Delta &= \frac{C(S+h) - C(S-k)}{h+k} \\
\Gamma &= 2\left(\frac{C(S+h) - C(S)}{h} - \frac{C(S) - C(S-k)}{k}\right) / (h+k)
\end{aligned}
\tag{2}
$$

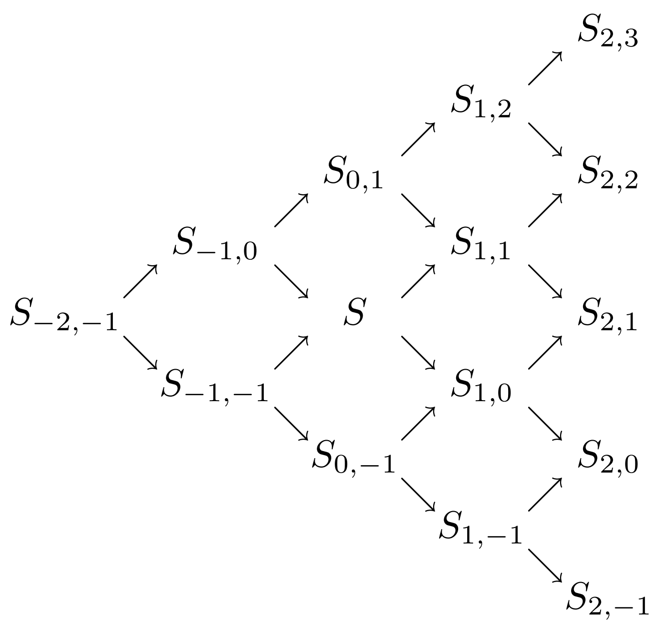


Figure 2. Extended Binomial Tree

A well known alternative method to calculate Δ and Γ is to extend the original tree as in Figure 2, where we still have $S_{n,j} = Su^j d^{n-j}$. If we calculate the option value in all the points of the extended tree, we can approximate the derivatives by

$$\begin{aligned}\Delta &= \frac{C(S_{0,1}) - C(S_{0,-1})}{S_{0,1} - S_{0,-1}} \\ \Gamma &= 2\left(\frac{C(S_{0,1}) - C(S)}{S_{0,1} - S} - \frac{C(S) - C(S_{0,-1})}{S - S_{0,-1}}\right) / (S_{0,1} - S_{0,-1})\end{aligned} \tag{3}$$

The alternative method is faster since one only has to evaluate one tree with $(n+3)(n+2)/2$ nodes instead of three trees with $(n+1)n/2$ nodes each. Furthermore, the alternative method can be interpreted as a specific case of the straightforward numerical differentiation method when we put $h = S_{0,1} - S$ and $k = S - S_{0,-1}$ in (2). However, this might indicate a weakness of the alternative method since derivatives are defined as the limits of expression (2) where h and k go to zero. In the extended tree one uses specific values for h and k and one wonders whether these are close enough to zero.

Figure 3. Comparison of Δ's

Figure 4. Comparison of Γ's

2. Numerical Results

Both methods are illustrated with the following example: Assume a stock price with a volatility of 25%, a time to maturity of the option of 1 year, while the exercise price is 100. The riskless interest rate is 9%. We vary the stock price. The number of periods in the tree is 50.

Figure 3 gives three different graphs of delta for different stock prices. First, the analytical values of delta from the Black-Scholes formula are plotted. Second, the deltas calculated by the numerical differentiation formula with $h = 0.01$ and finally, the values of delta from the extended binomial tree are plotted. The extended tree method, in contrast to our doubts, gives very good approximations of the true Black-Scholes deltas and certainly much better approximations than the numerical differentiation method.

Figure 4 gives graphs for gamma calculated according to the Black-Scholes formula, numerical differentiation formula with $k = 0.01$ and the extended tree formula, respectively. Once again, the extended tree method leads to very good approximations of the Black-Scholes Γ, while the numerical differentiation method gives approximations that do not seem to make any sense.

3. Why is Numerical Differentiation Inaccurate?

One wonders why the extended tree method is so much more accurate than numerical differentiation. Were h and k in the numerical differentiation method not close enough to zero? No, we show that they might even be too close to zero.

It is well-known that the binomial model gives a very good approximation of the Black-Scholes values. The binomial value is given by the following expression

$$C(S) = r^{-N} \sum_j \binom{N}{j} \pi^j (1-\pi)^{N-j} \max\{Su^j d^{N-j} - K, 0\} \tag{4}$$

where π is the risk-neutral probability. $C(S)$ is a piecewise-linear function of S. Figure 5 gives a graph of a piecewise linear function for $N = 2$. In the numerical differentiation method the derivatives of the piecewise linear function are calculated. It is clear that the function is not differentiable in the kinks of the graph, while the first derivative is constant

Figure 5. Piecewise linear function

between two kinks. This agrees exactly with the deltas given in Figure 3 for the numerical differentiation method. Furthermore, the second derivatives of the piecewise linear function are all zero. However, if by coincidence one uses points that are in different parts of the piecewise linear function in formula , one finds very large values for the derivatives. Of course, these values make no sense.

The above illustrates a well-known fact from numerical mathematics that a good approximation of a function does not imply that the derivatives of the approximating function are reliable approximations of the true derivatives.

In the extended tree model the delta is exactly equal to the hedge ratio of the replication strategy of the option. It is the value of the replication strategy that gives the binomial option its value. In the extended tree method one only compares points at different linear parts in the graph for calculating the derivatives. Hence, one is not hindered by the piecewise linearity.

4. Conclusions and Remarks

It can be concluded that the extended binomial tree method is not only faster but also more reliable than the numerical differentiation method to calculate delta and gamma of options. Although only European options were considered, the same conclusions can be drawn for American-style options. The binomial approximations for these options are piecewise linear as well. There is no formula like (4) but one can easily prove the piecewise linearity by a backward induction argument.

The derivative of the option price with respect to time theta (Θ) can be calculated from the option price, delta and gamma by using the Black-Scholes partial differential equation. Hence, good approximations of delta and gamma immediately give accurate approximations of theta. Alternatively, one can calculate theta by comparing $C(S_{-2,-1})$ and $C(S)$. For the derivatives with respect to the interest rate (rho) and volatility (vega) one has to use numerical differentiation methods. However, the binomial option price formula is not a piecewise linear function in these two variables and so we do not expect the same problems with numerical differentiation.

Sometimes, a variant of the extended binomial tree is used where one does not extend the tree but calculates delta from the first period option prices and gamma from the second period option prices. This method is less accurate than the extended tree method especially for a small number of periods in the tree, but certainly gives better results than the numerical differentiation method, since it also circumvents the pitfall of the piecewise linearity of the binomial pricing model.

9

Optimal Optioned Portfolios with Confidence Limits on Shortfall Constraints

OFTEN a portfolio manager faces the problem of composing a portfolio that behaves in line with certain risk-return constraints specified by an investor's preference schemes.[1] Options can be powerful instruments in obtaining the desired portfolio return distribution. In order to use options effectively, a number of conditions must be met. First, the manager must be able to track the changes in the distribution due to adding options to a stock portfolio. Second, he must have a clear understanding of the risk preferences of the investors. Finally, he must be able to select the best portfolio of stocks and options that matches with these preferences.

In this chapter, attention is focused on the work of a portfolio manager who has to construct a portfolio of stocks and options for a client or a group of investors. It is expected that the client will stick to the same portfolio over some specified time interval, after which an evaluation of the portfolio's performance will be made. In other words, the manager does not revise the portfolio continuously and will not use continuous trading in stocks and riskless loans to synthetically create options. The manager essentially follows a buy-and-hold strategy. In this case, options add extra possibilities to influence the final return distribution of a stock portfolio that cannot be created by only investing in stocks.

It is virtually impossible to describe the exact distribution of an

[1] This chapter is based on the paper with the same title by A. Pelsser and T. Vorst, which was published in Advances in Quantitative Analysis of Finance and Accounting (1995).

optioned portfolio with more than four stocks. Bookstaber and Clarke (1983) have described a method to approximate the distribution of such an optioned portfolio. This paper uses more or less the same method, although the underlying distributions of future stock prices will differ. The main difference is that it is assumed that future stock prices can be described by log-normal distributions and not by normal distributions. It is common practice in financial markets to use the Black-Scholes formula for the pricing of options. However, one of the assumptions underlying this formula is that future stock prices are log-normally distributed. Hence, in contrast with Bookstaber and Clarke (1983), this paper's assumptions on stock prices and option prices are mutually consistent. Also, these assumptions do not give rise to negative stock prices with a positive probability. Although the reality of log-normal distributions in describing stock prices might be questionable, it is even harder to get full agreement on a model for the risk preferences of a group of investors in a fund. It is often assumed that the risk preferences can be modeled by a Von Neumann-Morgenstern utility function (see, e.g. Ingersoll (1987)), but even then it is hard to specify the correct form of this function for the specific group. In order to avoid the specification of a Von Neumann-Morgenstern utility function and at the same time to take the risk aversion of the clients into account, this paper will assume that the portfolio manager is able to specify confidence limits on shortfall constraints for the client. The confidence limit approach has been described for stock portfolios by Leibowitz and Henriksson (1989). In their approach, it is assumed that the client can specify some shortfall constraints and, for each of the constraints, the highest probability the client will accept that the portfolio underperforms the shortfall constraint. An investor might, for example, accept a probability of at most 5% that the portfolio ends below 90% of today's value and a probability of at most 15% that the portfolio ends below today's value.

Leibowitz and Henriksson (1989) give an example of a fund manager whose performance is measured by comparison with an index fund. It is clear that this kind of performance measurement might also be relevant for the manager of the optioned portfolio in mind here. Thus, it will be assumed that all the knowledge of the risk preferences can be incorporated into confidence limits on shortfall constraints.

Having described the way the true distribution of optioned portfolios and the structure of the risk preferences are approximated, the paper comes to the final point of constructing the best optioned portfolio. It will be assumed that the client expects the manager to choose the portfolio with the highest expected return under the condition that the confidence limits on shortfall constraints are satisfied. With this definition of the best portfolio, linear programming techniques can be used to find the optimal portfolio. Although most investors use expected return and variance as yardsticks for measuring portfolio performance, here only return value, and not variance, will be used, as variance is no longer a sufficient statistic for optioned portfolios, as it is for stock portfolios. Also higher moments matter, and above that, the confidence limit method bounds the variance if the client correctly states the constraints. The method focuses on the distribution of the final payoffs and is, in this respect, similar to the payoff distribution pricing model (PDPM) described in Dybvig (1988a, 1988b). However, getting a good approximation of the final payoff distribution will mean imposing restrictions on the feasible portfolio compositions. These restrictions imply that the portfolio value will be an increasing function of the value of the market portfolio. Hence, the states of the world in which one has a high portfolio value coincides with a bullish stock market. In this respect the model is different from Dybvig's.

The chapter is organised as follows. Section 1 develops the model for approximating the distribution of optioned portfolios. Section 2 discusses how the restrictions on the preferences can be viewed as linear restrictions in a mathematical programming problem. Section 3 solves some specific portfolio optimization problems and Section 4 concludes.

1. Approximation of the Return Distribution

To describe the distribution of the future portfolio value, we first have to model the return distribution of the stocks and options. As explained in the introduction, log-normal distributions are used for future stock prices, since these are in line with the Black-Scholes formula. As in Bookstaber and Clarke (1983), a single market index or market model is used to describe the relation between the returns and (co)variances of the different stocks. This implies the use of the CCAPM, as described in Merton (1992, Ch. 15), which is consistent with Black-Scholes option pricing theory and comes with log-normal distributions for future stock prices. This model is briefly described below.

In the CCAPM there exists a market portfolio M, where the price process S_{M} can be described by a geometric Brownian motion

$$dS_{\mathrm{M}} = \mu_{\mathrm{M}} S_{\mathrm{M}}\, dt + \sigma_{\mathrm{M}} S_{\mathrm{M}}\, dW_{\mathrm{M}}, \tag{1}$$

where the constants μ_{M} and σ_{M} are the instantaneous drift and volatility of the market portfolio and W_{M} is a Wiener process. For an individual stock price S_i, the following price process is specified

$$dS_i = \mu_i S_i\, dt + \eta_{i\mathrm{M}} S_i\, dW_{\mathrm{M}} + \eta_{ii} S_i\, dW_i, \tag{2}$$

where the constant μ_i is the instantaneous drift of S_i, W_i are independent (also of W_{M}) Wiener processes and $\eta_{i\mathrm{M}}$ and η_{ii} are constants.

From (2), it can be seen that each S_i is described by a geometric Brownian motion with instantaneous variance equal to $\eta_{i\mathrm{M}}^2 + \eta_{ii}^2$. Hence, the volatility of the ith stock, as used in the Black-Scholes formula, is given by $\sqrt{\eta_{i\mathrm{M}}^2 + \eta_{ii}^2}$, while the instantaneous covariance of S_i and S_j is $\eta_{i\mathrm{M}}\eta_{j\mathrm{M}}$.

The CCAPM now gives the following equilibrium relation between returns and (co-) variances (Merton, 1990, Theorem 15.1)

$$\mu_i - r = \frac{\eta_{i\mathrm{M}}\sigma_{\mathrm{M}}}{\sigma_{\mathrm{M}}^2}(\mu_{\mathrm{M}} - r), \tag{3}$$

where r is the risk-free interest rate. The term $\eta_{i\mathrm{M}}\sigma_{\mathrm{M}}$ represents the covariance between S_i and S_{M}.

Equations (1), (2) and (3) specify the continuous-time single index or market model. In the sequel, we only use that $S_{\mathrm{M}}(T)$ and the individual $S_i(T)$ are log-normally distributed[2] and the equilibrium relation (3). It is not necessarily true that the S_i, $i = 1, \ldots, n$ form a basis of the marketed assets, as is usually assumed in the derivation of the CCAPM. They are just stocks that the portfolio manager considers for investment.

Let T be the client's horizon date. Since the portfolio manager does not change the portfolio before time T, the distribution of an optioned

[2] The market portfolio is constantly rebalanced. Hence, it is not a buy-and-hold portfolio, in which case it could not have been a log-normal distribution itself, as the sum of log-normal distributions. Due to the constant rebalancing, this market portfolio is log-normally distributed. (See Merton, 1990, Theorem 15.1)

portfolio at time T will be determined here. As already remarked in the introduction, it is impossible to explicitly calculate the complete distribution of an optioned portfolio if it contains more than four stocks. Therefore, an approximation procedure similar to that of Bookstaber and Clarke (1883) will be used. In this chapter we will only describe the procedure in case all options expire at the horizon date T. However, it is straightforward to apply the approximation procedure to options with longer maturities.

For every possible value $\bar{S}_{\mathrm{M}}(T)$ of the index, the conditional expected value of the optioned portfolio Q is calculated. This value will be denoted by $E\big(Q(T) \mid \bar{S}_{\mathrm{M}}(T)\big)$, which can be found by calculating and summing the conditional expectations of the individual securities, i.e. stocks, calls and puts, in the portfolio. This can be done as follows. Equations (1) and (2) imply that

$$\log\left(\frac{S_{\mathrm{M}}(T)}{S_{\mathrm{M}}(0)}\right) = (\mu_{\mathrm{M}} - \tfrac{1}{2}\sigma_{\mathrm{M}}^2)T + \sigma_{\mathrm{M}}\sqrt{T}X_{\mathrm{M}}, \tag{4}$$

where X_{M} is a standard normal random variable and

$$\log\left(\frac{S_i(T)}{S_i(0)}\right) = (\mu_i - \tfrac{1}{2}\eta_{i\mathrm{M}}^2 - \tfrac{1}{2}\eta_{ii}^2)T + \eta_{i\mathrm{M}}\sqrt{T}X_{\mathrm{M}} + \eta_{ii}\sqrt{T}X_i, \tag{5}$$

where the X_i are mutually independent standard normal random variables (also independent of X_{M}). With each realization $\bar{S}_{\mathrm{M}}(T)$ of the random variable $S_{\mathrm{M}}(T)$, there corresponds (according to (4)) a value $\bar{X}_{\mathrm{M}}$ for the random variable X_{M}. Hence, to calculate $E\big(S_i(T) \mid \bar{S}_{\mathrm{M}}(T)\big)$, set X_{M} equal to $\bar{X}_{\mathrm{M}}$ in (5) and calculate the expected value. This leads to

$$E\big(S_i(T) \mid \bar{S}_{\mathrm{M}}(T)\big) = S_i(0)e^{v_i(\bar{X}_{\mathrm{M}})}, \tag{6}$$

with

$$v_i(\bar{X}_{\mathrm{M}}) = (\mu_i - \tfrac{1}{2}\eta_{i\mathrm{M}}^2)T + \eta_{i\mathrm{M}}\sqrt{T}\bar{X}_{\mathrm{M}}. \tag{7}$$

For a call option with exercise price K, the conditional expected value $E\big(C_i(T) \mid \bar{S}_{\mathrm{M}}(T)\big)$ is given by the following formula, which can be deduced in a way similar to the Black-Scholes formula

$$E\big(C_i(T) \mid \bar{S}_{\mathrm{M}}(T)\big) = S_i(0)e^{v_i(\bar{X}_{\mathrm{M}})}N(d_1) - KN(d_2), \tag{8}$$

with

$$d_1 = \frac{\log(S_i(0)/K) + v_i(\bar{X}_{\mathrm{M}}) + \frac{1}{2}\eta_{ii}^2 T}{\eta_{ii}\sqrt{T}}$$
$$d_2 = d_1 - \eta_{ii}\sqrt{T} \tag{9}$$

For a put option the conditional expected value is given by

$$E\big(\,P_i(T)\;\big|\;\bar{S}_{\mathrm{M}}(T)\,\big) = KN(-d_2) - S_i(0)e^{v_i(\bar{X}_{\mathrm{M}})}N(-d_1), \tag{10}$$

Notice that (8) and (10) are not option prices, but the future expected values of the options. Hence, the expected growth of the stock was used instead of the risk-neutral rate, and the future expected value was not discounted to get today's value.

Let $K_{i1} < K_{i2} < \cdots < K_{in_i}$ be the possible exercise prices of calls or puts in the ith stock. Assume the portfolio consists of α_i stocks with price S_i, β_{ij} calls on the ith stock with exercise price equal to K_{ij}, and γ_{ij} puts on the ith stock with exercise price K_{ij}. The conditional expected value of the portfolio is then given by

$$E\big(\,Q(T)\;\big|\;\bar{S}_{\mathrm{M}}(T)\,\big) = \sum_{i=1}^{n}\alpha_i E\big(\,S_i(T)\;\big|\;\bar{S}_{\mathrm{M}}(T)\,\big)$$
$$+\sum_{i=1}^{n}\sum_{j=1}^{n_i}\Big[\beta_{ij}E\big(\,C_{ij}(T)\;\big|\;\bar{S}_{\mathrm{M}}(T)\,\big) + \gamma_{ij}E\big(\,P_{ij}(T)\;\big|\;\bar{S}_{\mathrm{M}}(T)\,\big)\Big], \tag{11}$$

where the index j in C_{ij} and P_{ij} refers to the exercise price K_{ij} of the options.

Up to now no restrictions have been placed on the values of α_i, β_{ij} and γ_{ij}. For example, γ_{ij} might be negative, which would mean that the portfolio manager has written put options on the ith stock with exercise price K_{ij}. However, it will have to be assumed for the moment that $E\big(\,Q(T)\;\big|\;\bar{S}_{\mathrm{M}}(T)\,\big)$ is an increasing function of $\bar{S}_{\mathrm{M}}(T)$. For the purpose at hand, constructing a portfolio that is protected with high probability against a strong depreciation, this is not unrealistic. Investors in such a fund are unlikely to speculate against the market as a whole. Also, the portfolio manager that Leibowitz and Henriksson (1989) have in mind, whose performance is measured against a benchmark portfolio, is unlikely to speculate against market movements.

Sufficient conditions for making $E\big(\,Q(T)\;\big|\;\bar{S}_{\mathrm{M}}(T)\,\big)$ an increasing function of $\bar{S}_{\mathrm{M}}(T)$ are

$$\eta_{i\mathrm{M}} > 0 \qquad \forall i \tag{12}$$

$$\alpha_i + \sum_{j=1}^{m} \beta_{ij} - \sum_{j=m+1}^{n_i} \gamma_{ij} > 0 \qquad \forall i \quad \forall m = 0, \ldots, n_i. \tag{13}$$

The first condition implies considering only those stocks that are positively correlated with the market. The second condition implies that, at maturity, the value of the part of the portfolio consisting of the ith stock and the options written on that stock is an increasing function of the stock price over every interval $[K_{i,m}, K_{i,m+1}]$. We can weaken these conditions somewhat, by allowing for stocks that are negatively correlated with the market and for which the inequality sign in (13) is reversed.

Let $g(.)$ be the inverse function of $E\big(\,Q(T)\;\big|\;.\,\big)$. The distribution function F of the value of the portfolio Q at time T is approximated by

$$F(x) = \Pr[Q(T) \le x] \approx \Pr\big[E\big(\,Q(T)\;\big|\;S_{\mathrm{M}}(T)\,\big) \le x\big] = \Pr[S_{\mathrm{M}}(T) \le g(x)] = G\big(g(x)\big) \tag{14}$$

where G is the distribution function of $S_{\mathrm{M}}(T)$. Formulas (11) and (14) jointly determine the approximation of the distribution function of the value of the portfolio Q at time T. The approximation is accurate if the aggregated influence of the sources of nonsystematic risk W_i is small compared to the influence of the systematic risk W_{M} on the value of the portfolio Q. In the limit, when the number of stocks becomes large and the nonsystematic risk cancels out, the approximation will be equal to the exact distribution. Bookstaber and Clarke (1983) tested the approximation for the distribution for portfolios of five different stocks by comparing it with the distribution of $Q(T)$ generated by Monte Carlo simulation. The results are described in their paper and are very satisfying. Although our model differs from theirs with respect to the model describing the stock price behaviour (a continuous vs. a discrete single-index model) the approximation algorithm does equally well, since the procedure is essentially the same. This was confirmed by Monte Carlo tests similar to those of Bookstaber and Clarke.

In using options with a maturity date past the horizon date T, one can use enhanced versions of (8) and (10), which calculate the expected Black-Scholes value of the options at time T. It is fairly straightforward to derive these formulas.

2. The Portfolio Optimization Problem

It is assumed that the portfolio manager has a fixed budget $Q(0)$ to invest and that the return is measured against this budget. Of course, he might want to invest part of the budget in riskless bonds that mature at time T. Thus, assume that security number n, with value S_n, is a riskless bond with risk-free rate $r > 0$, for which $\eta_{n\mathrm{M}} = \eta_{nn} = 0$. In this case, (3) reduces to $\mu_n = r$. It is also assumed that no options are traded on asset n, since it is riskless. The budget restriction reads

$$\sum_{i=1}^{n} \alpha_i S_i(0) + \sum_{i=1}^{n-1} \sum_{j=1}^{n_i} \big[\beta_{ij} C_{ij}(0) + \gamma_{ij} P_{ij}(0)\big] \leq Q(0), \tag{15}$$

where $S_i(0)$, $C_{ij}(0)$ and $P_{ij}(0)$ are today's prices of the stocks, calls and puts. Because $r > 0$, it is in the interest of the investors that the whole budget $Q(0)$ is invested. Thus, in the optimal solution, the inequality will actually be an equality.

We now turn to incorporating the confidence limits on shortfall constraints in the portfolio optimization problem. Suppose that the fund manager, according to the wishes of the investors in the fund, wants to set a shortfall constraint at a return of -5%, and is willing to accept a probability of at most 10% that the fund underperforms this constraint, i.e.

$$\Pr[Q(T) \geq 0.95 Q(0)] \geq 0.90. \tag{16}$$

This constraint has to be translated into a constraint on the α_i, β_{ij} and γ_{ij}, the decision variables in the portfolio composition problem.

Choose $\bar{y}(T)$ such that $G\big(\bar{y}(T)\big) = 0.10$. Recall that G is the cumulative distribution function of $S_{\mathrm{M}}(T)$. If $E\big(\,Q(T) \bigm| \bar{y}(T)\,\big) \geq 0.95Q(0)$, then $\bar{y}(T) \geq g\big(0.95Q(0)\big)$ and we obtain (using our approximation for the distribution of Q)

$$\begin{aligned}
\Pr[Q(T) \geq 0.95Q(0)] &\approx \Pr\big[E\big(\,Q(T) \bigm| S_{\mathrm{M}}(T)\,\big) \geq 0.95Q(0)\big] \\
&= \Pr\big[S_{\mathrm{M}}(T) \geq g\big(0.95Q(0)\big)\big] \\
&\geq \Pr[S_{\mathrm{M}}(T) \geq \bar{y}(T)] = 0.90
\end{aligned} \tag{17}$$

and vice versa. Therefore, (16) can be rewritten as

$$E\big(\,Q(T) \bigm| \bar{y}(T)\,\big) = \sum_{i=1}^{n} \alpha_i E\big(\,S_i(T) \bigm| \bar{y}(T)\,\big)$$

$$+\sum_{i=1}^{n-1}\sum_{j=1}^{n_i}\left[\beta_{ij}E\big(\,C_{ij}(T)\;\big|\;\bar{y}(T)\,\big)+\gamma_{ij}E\big(\,P_{ij}(T)\;\big|\;\bar{y}(T)\,\big)\right]\geq 0.95Q(0). \quad (18)$$

Since all terms, except for the α, β and γ can be calculated in advance, (18) is a linear restriction in α, β and γ. Using this procedure, the portfolio manager can add as many restrictions as necessary to reflect the risk preferences of the investors.

Now that the confidence limits on shortfall constraints have been translated to linear restrictions of the form (18), we turn to the objective function that has to be maximized. It is assumed that the investor prefers the portfolio with the highest expected returns that fulfils the restrictions. Hence, the portfolio manager wants to maximize

$$E\big(Q(T)\big)=\sum_{i=1}^{n}\alpha_i E\big(S_i(T)\big)+\sum_{i=1}^{n-1}\sum_{j=1}^{n_i}\beta_{ij}E\big(C_{ij}(T)\big)+\gamma_{ij}E\big(P_{ij}(T)\big), \quad (19)$$

which is a linear function in the decision variables α, β and γ. Consequently, the portfolio optimization problem can be rewritten as a linear programming problem.

Depending on the number of different stocks, of different puts and calls on these stocks, and of shortfall constraints, the optimization problem can be solved more or less straightforwardly. Of course, the optimization routine used to solve the problem might reveal that there are no feasible solutions. This might be caused by very strong confidence limits on some shortfall constraints. For example, if the risk-free rate over the horizon date is 7% and the fund manager faces a constraint like

$$\Pr[Q(T)\geq 1.10Q(0)]\geq 0.95, \quad (20)$$

it might be very hard to find a portfolio that fulfils this constraint. Furthermore, it might also happen that some combination of restrictions lead to infeasibility of the optimization problem. Thus, although it was stated before that the fund manager can add as many restrictions as is necessary to reflect the risk preferences of the investors, he should keep reality in mind. A good optimization routine indicates which restrictions leads to infeasibility.

Although the portfolio optimization problem was described in fairly general terms, the method used above can also be applied to cases in which the fund manager does not have perfect freedom in choosing stocks, calls and puts. An example might be a manager confronted with a given portfolio of stocks and asked to reshape the return distribution of the portfolio by buying and/or writing options, to reflect the risk characteristics of the investors in the portfolio. In this case, he fixes all the α, leaving the β and γ as decision variables. Another example might be a client who does not want to write puts. In this case, the extra restrictions will be $\gamma_{ij} \geq 0$. In general, there might be upper and/or lower bounds on the number of options or stocks in the portfolio.

If the use of options is desired to reshape the return distribution of the portfolio, it makes sense to use more than two different shortfall constraint confidence limits to model the risk preferences of the investors. This is in contrast with the application described by Leibowitz and Henriksson (1989), where it is assumed that the returns of portfolios are normally distributed. In that case the return distribution can be completely specified by the mean and the variance; there will thus be only two restrictions that are binding, while the others will be redundant. Since the return distribution of a portfolio of stock options cannot be completely described by any finite number of its moments, it makes perfect sense to use several shortfall constraints on confidence limits.

Finally, it must be pointed out why it is sometimes necessary to use some extra restrictions. Suppose that the fund manager uses only restriction (16). Assume that there exists a put option $\bar{P}$, which is priced such that

$$\bar{P}(0) \geq e^{-rT} E\big(\bar{P}(T)\big), \tag{21}$$

and the probability that the put option will be exercised is only 7.5%. In that case, a solution is to write infinitely many of these puts and invest the proceeds in the riskless asset S_n. This gives an infinite value of the objective function, due to condition (21). Restriction (16) is satisfied, since the probability that the option is exercised is only 7.5%, which is well below the confidence limit of 10%. Of course, nobody is interested in this portfolio, because in case of a loss it will be infinitely large, even if the chances of a loss are small. If the manager adds the constraint

$$\Pr[Q(T) \geq 0.90 Q(0)] \geq 0.95, \tag{22}$$

the portfolio described before is no longer allowed. The same effect can be achieved by putting upper and lower bounds on the number of individual options in the portfolio, as described before.

Table I: Stocks in portfolio

i	$S_i(0)$	$\eta_{i\mathrm{M}}$	η_{ii}
1	100	0.4	0.05
2	100	0.5	0.05
3	100	0.5	0.10
4	100	0.6	0.05
5	100	0.6	0.10

3. An Example of Portfolio Optimization

This section illustrates the methodology described in the previous sections with an example. Five different stocks and no bonds will be used; their characteristics are given in Table I. Set $T = 0.5$, $\eta_{\mathrm{M}} = 0.10$, $r = 0.05$ and $\sigma_{\mathrm{M}} = 0.25$. From these and equation (3) the η_i can be deduced. Of course, it is not a restriction to assume that all stocks have the same price today, since the amount invested in a stock can be considered as a decision variable. Furthermore, we assume that for each stock there are traded call and put options with exercise prices 90, 95 and 110 (referred to as C90, C95 and C110 for the calls, and P90, P95 and P110 for the puts, for the remainder of this section). The maturity date for all options coincides with the horizon date T. Finally, we assume that the budget $Q(0)$ is equal to 1000. The portfolio manager uses the following confidence limits on shortfall constraints

$$\begin{aligned}
&\Pr[Q(T) \geq 0.80Q(0)] \geq 0.99\\
&\Pr[Q(T) \geq 0.90Q(0)] \geq 0.95\\
&\Pr[Q(T) \geq 0.95Q(0)] \geq 0.90\\
&\Pr[Q(T) \geq 1.00Q(0)] \geq 0.75\\
&\Pr[Q(T) \geq 1.05Q(0)] \geq 0.55.
\end{aligned} \tag{23}$$

The optimal solution of the problem under these shortfall constraints, the budget constraint, and the monotonicity restrictions (13), is reported in Table II. In the optimal portfolio, there is a long position in stocks 1 and 5. For stock 1, this is combined with a short position in C90 to profit from the option premium. For stock 5, there is really no long position, since the short C110 and the long P110, together with the stock, amount to a long position in bonds via put-call parity. It is also interesting to see that a bullish vertical spread is created with the options on stock 2.

Table II: Holdings of different stocks and options in optimal portfolio

	1	2	3	4	5
Stocks	3.755	0.000	0.000	0.000	6.200
C90	−3.755	6.815	0.000	0.000	0.000
C95	0.000	−6.815	0.000	0.000	0.000
C110	0.000	0.000	0.000	0.570	−6.200
P90	0.000	0.000	0.000	0.000	0.000
P95	0.000	0.000	0.000	0.000	0.000
P110	0.000	0.000	0.000	0.000	6.200

$E(Q(T)) = 1033.80$

Table III: Holdings of different options in optimal portfolio with fixed number of stocks

	1	2	3	4	5
Stocks	2.000	2.000	2.000	2.000	2.000
C90	−2.000	4.815	−2.000	−2.000	4.001
C95	0.000	−6.815	0.000	0.000	0.000
C110	0.000	0.000	0.000	0.570	−6.001
P90	−1.755	2.000	2.000	2.000	−4.001
P95	0.000	0.000	0.000	0.000	0.000
P110	0.000	0.000	0.000	0.000	6.001

$E(Q(T)) = 1033.80$

Figure 1: Final value distribution

To fully illustrate the power that options have in reshaping a portfolio's payoff distribution, we also solved a more restricted case of the problem just described. Suppose that the portfolio manager has already used his complete budget to buy stocks. Can he still modify the portfolio distribution by using puts and calls, even if he is not allowed to borrow extra money? The answer is yes, and is reported in Table III. It is clear from Table III that it is possible to reach the same expected value of the portfolio at T. Due to put-call parity, the portfolio manager is able to neutralize the "unwanted" stocks and still create the return distribution desired.

In Figure 1 we have plotted for the optimal portfolio from Table II the approximating distribution function of the final portfolio value (dashed line) and the simulated distribution function of the final portfolio value (solid line). The distribution functions are virtually identical. The simulated distribution value is an empirical distribution function made after 10,000 runs in a Monte Carlo simulation of the movements of the five underlying stocks. Although the portfolio consists of only five stocks, it is clear that the approximating distribution coincides quite well with the true final value distribution.

The vertical lines marked shortfall constraints in Figure 1 depict the confidence limits on shortfall constraints. For example, the line at 950 has a height of 0.1, indicating that the portfolio can have at most a probability of $1 - 0.9 = 0.1$ that its final value will be below 950. The simulated distribution function indeed satisfies all the constraints.

4. Conclusion

We used a continuous-time version of an algorithm of Bookstaber and Clarke to approximate the distribution of the returns of an optioned portfolio. With this approximation as a basis, it was shown how a portfolio manager can find an optimal optioned portfolio if he knows confidence limits on shortfall constraints. This method was illustrated by an example of an optimally optioned portfolio.

The methodology can be applied generally, since it allows the portfolio manager to use extra restrictions if he needs to take some institutional constraints into account.

It is also straightforward to incorporate proportional transaction costs into our methodology. This can be done by splitting all decision variables in two parts: one for a long and one for a short position and by adding positivity constraints on all the new decision variables. After this procedure it is possible to attach different prices to long and short positions in stocks and options, which reflects the transaction costs.

One of the disadvantages of the linear programming methodology is that the optimal solution does not yield integer numbers, but fractions of options that have to be bought or sold. Of course, this is impossible, and most option contracts even have lot sizes of more than 100 stocks. This problem might be solved in two ways. First it is possible to round off the optimal number of options to the closest number of real contracts. Since the portfolio distribution is only approximated by the methodology described above, this might not be unreasonable. Second, an integer programming methodology similar to the linear programming methodology could be used.

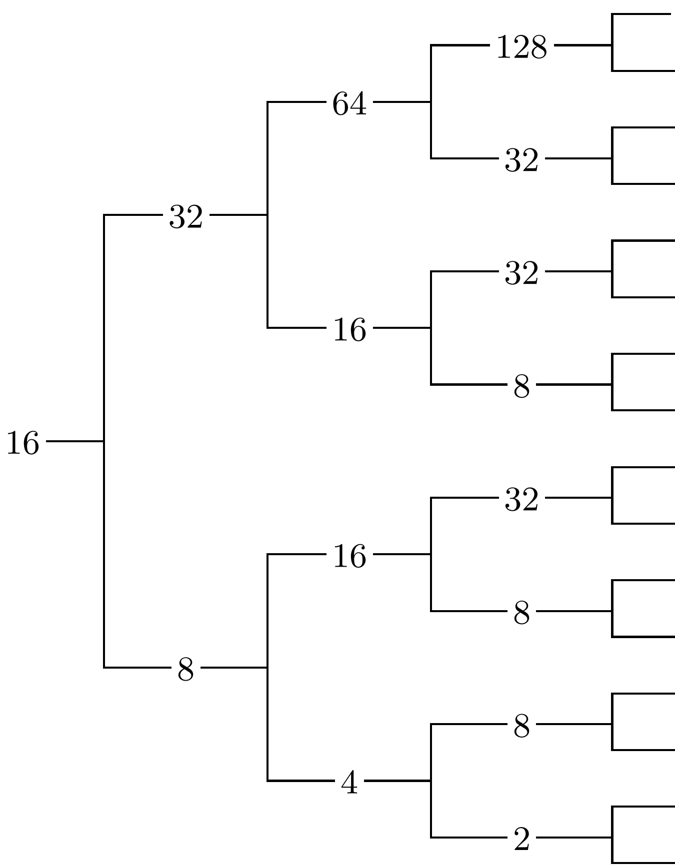
10

Transaction Costs and Efficiency of Portfolio Strategies

In an economy without frictions, such as transaction costs, options are redundant instruments, because their payoffs can be replicated by dynamic portfolio strategies with trading only in the underlying stock and riskless bonds.[1] Therefore, the price of an option is equal to the initial value of the portfolio strategy that replicates the option. In this setting the use of options does not create any new portfolio strategies. In a more realistic case where transaction costs are incurred on trades in stocks, Leland (1985) has shown that the transaction costs in replicating an option are substantial and that large differences exist between initial values of portfolios replicating long and short positions in options. In practice options are often priced according to the Black-Scholes formula which assumes no transaction costs. Hence, in a world with transaction costs, options might enhance the feasible portfolio strategies since they give payoff distributions that can only be obtained by a portfolio strategy in stocks and bonds at considerable transaction costs.

A more recent result for frictionless economies, established by Dybvig (1988b), is that popular dynamic portfolio strategies, such as stop-loss and lock-in, are inefficient. This means that there exists a portfolio strategy that gives the same final payoff distribution as the stop-loss strategy but has lower initial costs. The same holds for the lock-in strat-

[1] This chapter is based on the paper with the same title by A. Pelsser and T. Vorst, which is forthcoming in the European Journal of Operational Research. This research was sponsored by the Economics Research Foundation, which is part of the Netherlands Organisation for Scientific Research (NWO).



	State prob.	State price	State
256	1/16	1/81	ω_1
64	1/16	2/81	ω_2
64	1/16	2/81	ω_3
16	1/16	4/81	ω_4
64	1/16	2/81	ω_5
16	1/16	4/81	ω_6
16	1/16	4/81	ω_7
4	1/16	8/81	ω_8
64	1/16	2/81	ω_9
16	1/16	4/81	ω_{10}
16	1/16	4/81	ω_{11}
4	1/16	8/81	ω_{12}
16	1/16	4/81	ω_{13}
4	1/16	8/81	ω_{14}
4	1/16	8/81	ω_{15}
1	1/16	16/81	ω_{16}

$S_0 = 16,\ u = 2,\ d = \frac{1}{2},\ r = 1$

Figure 1: Expanded binomial tree

egy. These results depend heavily on the assumption of no transaction costs. The more efficient strategy generates more trading than the simple stop-loss strategy. Hence, in a world with transaction costs it is no longer obvious that the efficient strategies really command lower initial investments.

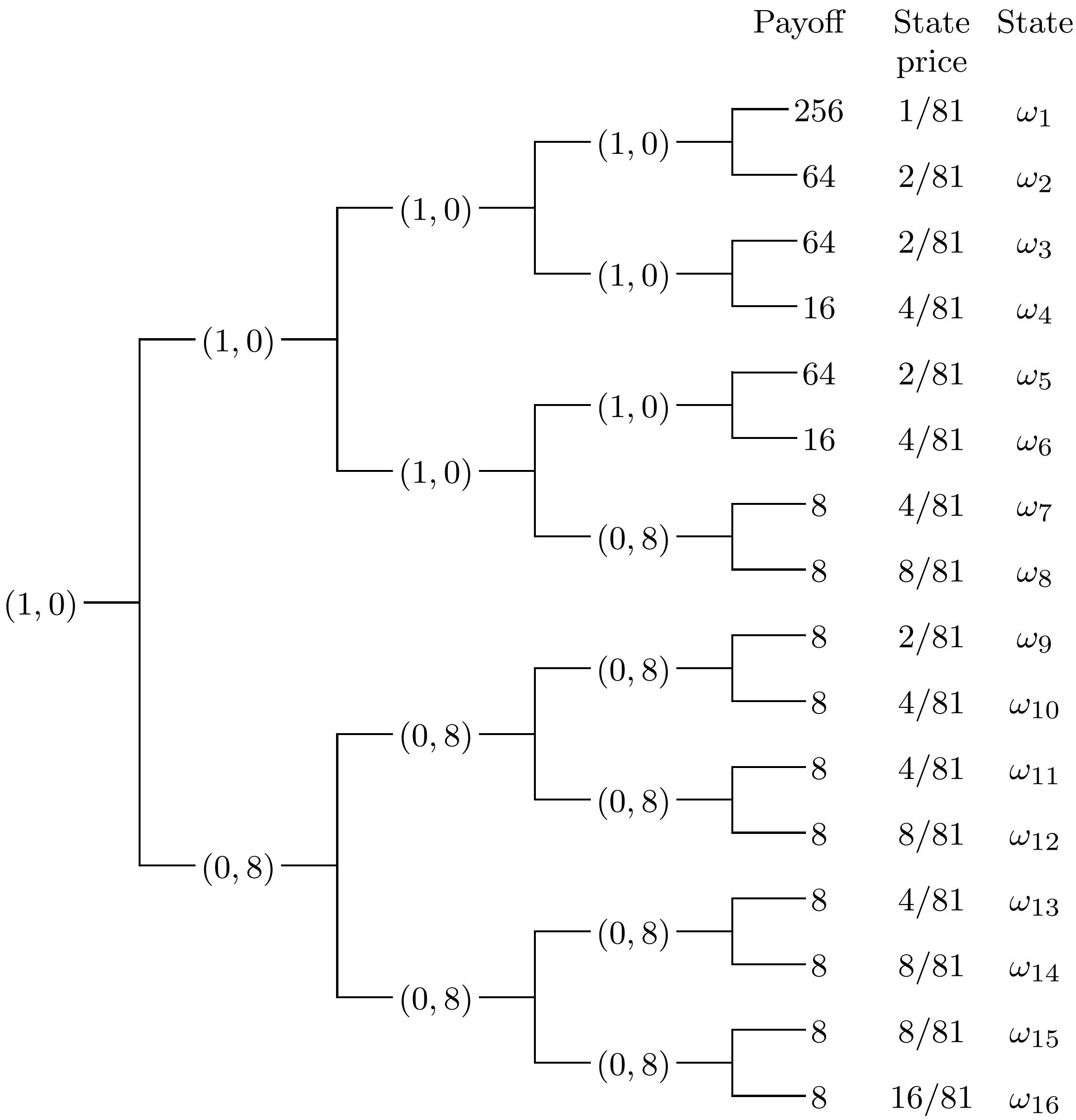
In this chapter we investigate the following two issues:

- Are simple dynamic portfolio strategies, such as stop-loss and lock-in, really inefficient if there are transaction costs on trades of stocks?
- Does the use of options in dynamic portfolio strategies enable investors to realize final value distributions at lower costs than would be possible with only trading in stocks and bonds?

To analyze these issues we use the Cox-Ross-Rubinstein (1979) binomial tree model for the stock prices. In the next section, we briefly describe the results of Dybvig (1988a,b) pertaining to the inefficiency of stop-loss and lock-in strategies. In Section 3 we introduce transaction costs, formulate the concepts of efficient and strongly efficient dynamic self-financing portfolio strategies and formulate optimization problems for dynamic portfolio strategies. Section 4 is devoted to the evaluation of efficient dynamic portfolio strategies when there are transaction costs, while Section 5 addresses the question whether options enhance the possibilities of investors to achieve certain payoff distributions at low costs. The last section concludes the paper and offers some directions for further research.

1. The Dybvig Model

In this section, we briefly outline the payoff distribution pricing model (PDPM) for the binomial asset pricing model as described in Dybvig (1988b). Figure 1 represents a specific (4-period) example of the well-known binomial model of stock returns introduced by Cox, Ross and Rubinstein (1979). The initial stock price S_0 is equal to 16, one plus the riskless interest rate is equal to one (hence, the riskless interest rate is zero) and the stock price doubles or halves each period, with probability 1/2. Since we are interested in both path-dependent trading strategies and transaction costs, we do not give the standard (recombined) binomial tree, but the expanded tree. In fact, the expanded tree properly shows the increasing sequence of σ-algebras that describes the information structure of the model.



Price of strategy: 16

$S_0 = 16,\ u = 2,\ d = \frac{1}{2},\ r = 1$

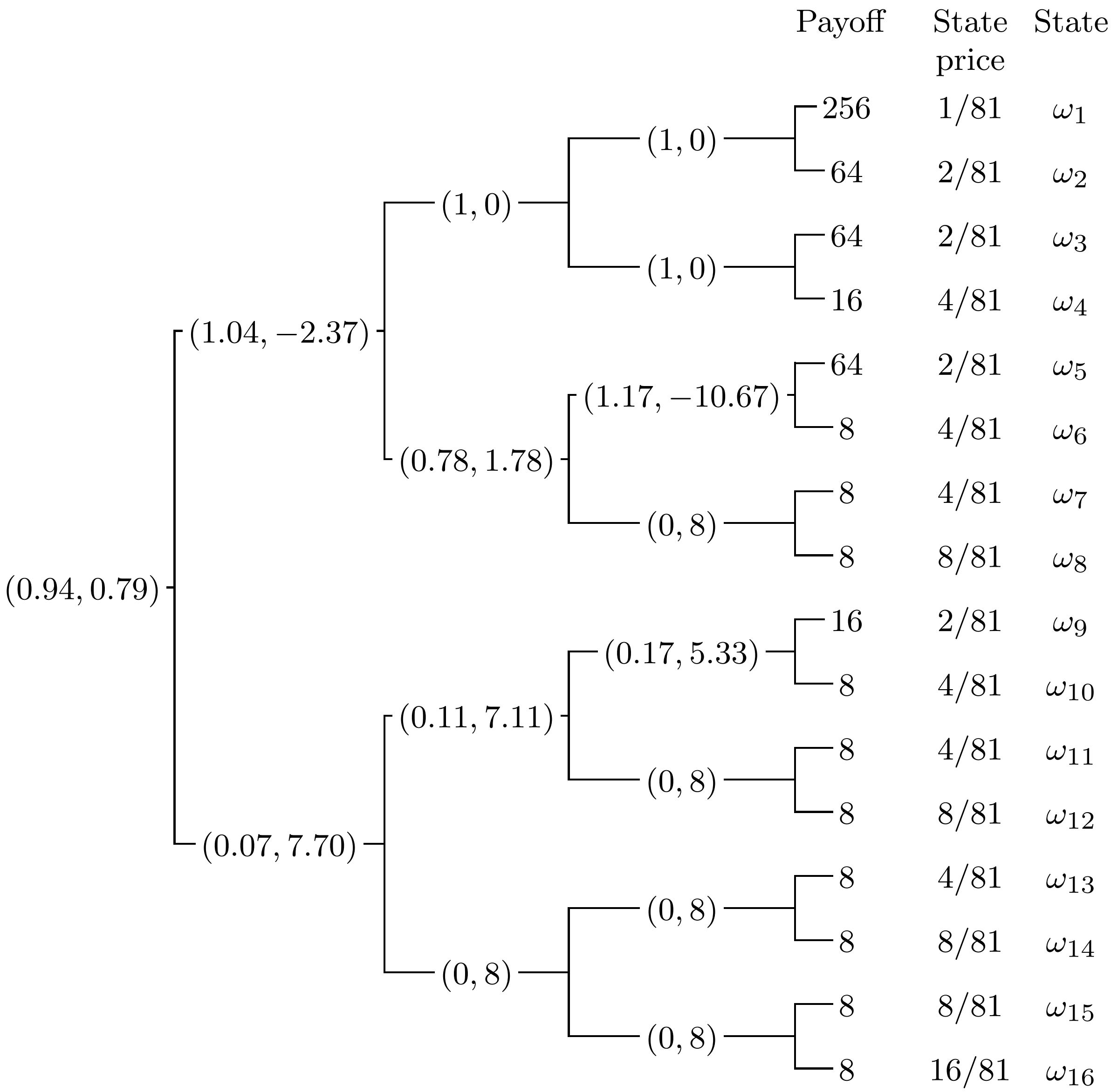
Portfolios: (# of stocks, amount in bonds)

Figure 2: Stop-loss strategy

The final states are numbered $\omega_1, \ldots, \omega_{16}$. Since the binomial model is complete, each contingent claim that gives payoffs in the final period can be duplicated by a self-financing trading strategy. If there are no arbitrage opportunities, the initial outlay for this strategy must be the price of the contingent claim. The price of a contingent claim that pays 1 in a given state and 0 in all other states is called the state price of the given state. For a given contingent claim its price is the sum over all states of the payment in that state times the state price. In fact, the state price is equal to the probability of the state under the unique equivalent martingale measure of the pricing model. Next, consider the stop-loss strategy where we initially buy one stock and sell the stock for bonds as soon as the stock price drops to 8. From there onward we deal solely with bonds. In Figure 2 this strategy is illustrated such that at each node in the tree the first number specifies the number of stocks in the portfolio and the second specifies the amount invested in bonds. Furthermore, the final value of the portfolio at maturity in each state is given. In Figure 3 an alternative to the stop-loss strategy is given. At each node, we again specify the number of stocks in the portfolio, and the amount invested in bonds. The alternative strategy is self-financing and has the same payoff distribution as the stop-loss strategy. Both strategies give a payoff of 8 with probability 10/16, a payoff of 16 with probability 2/16, a payoff of 64 with probability 3/16 and a payoff of 256 with probability 1/16. However, the initial outlay of the alternative strategy is only $15\frac{65}{81}$, i.e. $0.938 \times 16 = 15.01$ for the stock plus 0.79 for the bond. The reduction in price of the alternative strategy is due to the fact that the payoffs in states ω_6 and ω_9 are interchanged. The alternative strategy gives a higher payoff in a state with a lower state price and a lower payoff in a state with a higher price. The reduction in price is equal to the product of the difference in payoffs multiplied by the difference in state prices: $(16 - 8) \times (4/81 - 2/81) = 16/81$.

It is clear from the above example that an individual who maximizes expected utility of terminal wealth with a strictly increasing utility function prefers the alternative strategy to the stop-loss strategy. Dybvig (1988a) proves that the following properties of portfolio strategies are equivalent:

– The portfolio strategy c is chosen by some agent who has strictly increasing von Neuman-Morgenstern preferences over terminal wealth;



Price of strategy: 15.80

$S_0 = 16,\ u = 2,\ d = \frac{1}{2},\ r = 1$

Portfolios: (# of stocks, amount in bonds)

Figure 3: Alternative Stop-loss strategy

– The final portfolio value is non-increasing in the terminal state-price.[2]

Portfolio strategies that have the second property are called efficient, while other portfolio strategies will be called inefficient. If we compare the final portfolio values of the alternative strategy with the state price, it is clear that this strategy is efficient.

In Dybvig (1988b), the simple example above is extended to a binomial model with more trading intervals and non-zero interest rates. Dybvig calculates that the inefficiency of the stop-loss strategy might amount to 100 basispoints for an investment period of 1 year with a risk-free interest rate of 8%, a stock with an expected return of 16% and a volatility of 20% and a binomial tree of 360 trading intervals. This means there exists an alternative portfolio strategy with the same final payoff distribution as the stop-loss strategy that requires an initial outlay that is 1% lower.

In the binomial model, with upward and downward probability both equal to 1/2, the state price is a function of the final stock price. If two final states have the same stock price they will also have the same state price. If the expected return on the stock is higher then the risk-free return, then this function is strictly decreasing. Hence, a portfolio strategy is efficient if and only if the final portfolio value is a non-decreasing function of the final stock price. This implies that a portfolio that consists of only holding a put option is inefficient. However, a stock plus a put option on that stock is an efficient strategy. A call option itself is also an efficient strategy and it is clear that lock-in strategies, where the stock is sold if its price rises to a given level, are inefficient.

Nowadays, fund managers not only consider the risk and return of a portfolio strategy, but also compare the performance of their portfolio with a benchmark index (See e.g. Roll (1992)). If we replace the stock

[2] In fact, Dybvig's result is more general. The second property should be that the final portfolio value is non-increasing in the terminal state price density. The terminal state-price density can be obtained by dividing the state price by the state probability. The state price density is the likelihood ratio process or the risk adjustment process (see e.g. Dothan (1990) for an elaboration on the state price density). Since, in our case all terminal states have the same probability, state prices and state price densities only differ by a constant factor.

with the index in the model above, this implies that efficient strategies will have a final portfolio value that is increasing in the value of the index. This can be interpreted as efficient strategies being those strategies with the lowest tracking error given a final portfolio value distribution function.

2. The Influence of Transaction Costs

If we compare Figures 2 and 3, we see that although both strategies differ only in portfolio values in two states, the holdings in bonds and stocks differ substantially between the two strategies. For the stop-loss strategy we only have to sell the stock and buy the bond in some cases. For the alternative strategy the portfolio must be adjusted in each period, whatever happens with the stock value. Hence, it is not unlikely that the alternative strategy generates higher transaction costs. In this section we model these transaction costs. To a large extent we follow the notation from Bensaid, Lesne, Pagès and Scheinkman (1992).

There are $T+1$ dates, $t = 0, 1, \ldots, T$. The state space $\Omega = \{u, d\}^T$ is the set of paths in an expanded binomial tree, where u stands for an up-state movement and d for a down-state movement. It is assumed that at each step the probability of an up-state is 1/2. A path is denoted by $\omega = (\omega_1, \ldots, \omega_t, \ldots, \omega_T)$ with $\omega_i \in \{u, d\}$. $\omega^t = (\omega_1, \ldots, \omega_t)$ will denote a path up to time t and $\mathcal{F}_t$ is the σ-algebra of information up to time t. There is a risky asset called stock and a riskless asset called bond. The one-period return on the riskless bond is constant over time and states. One plus the riskless return will be denoted by r. The value of the stock at time t and state ω is denoted by $S_t(\omega)$, which is an adapted process. A dynamic portfolio strategy is defined by two $\mathcal{F}$-adapted processes Δ and B. Δ_t is the number of stocks that is held between date t and $t+1$ and B_t is the amount of dollars invested in the riskless bond during the same period. Transaction costs are due on trades in stocks and as in Boyle and Vorst (1992) they are proportional to the amount traded. Let k be the proportion of transaction costs. A dynamic portfolio strategy is called self-financing with transaction costs k (or shortly self-financing) if the following equation holds

$$\begin{aligned}\big(\Delta_{t-1}(\omega) - \Delta_t(\omega)\big)S_t(\omega) + rB_{t-1}(\omega) - B_t(\omega) &= \\ k\big|\Delta_{t-1}(\omega) - \Delta_t(\omega)\big|S_t(\omega) \qquad \forall t \le T-1,\ \forall \omega \in \Omega. &\quad (1)\end{aligned}$$

The final payoff V of a portfolio strategy is given by

$$V(\omega) = \Delta_{T-1} S_T(\omega) + r B_{T-1}(\omega), \tag{2}$$

which is a real-valued random variable on Ω. Let

$$G(\Delta, B)(x) = \Pr\big\{\omega \in \Omega \;\big|\; \Delta_{T-1} S_T(\omega) + r B_{T-1}(\omega) \le x\big\} \tag{3}$$

be the probability distribution function of this random variable. We call $G(\Delta, B)$ the payoff distribution function. Note that Δ_0 and B_0 do not depend on ω since (Δ, B) is adapted. $\Delta_0 S_0 + B_0$ is called the initial cost of the strategy.

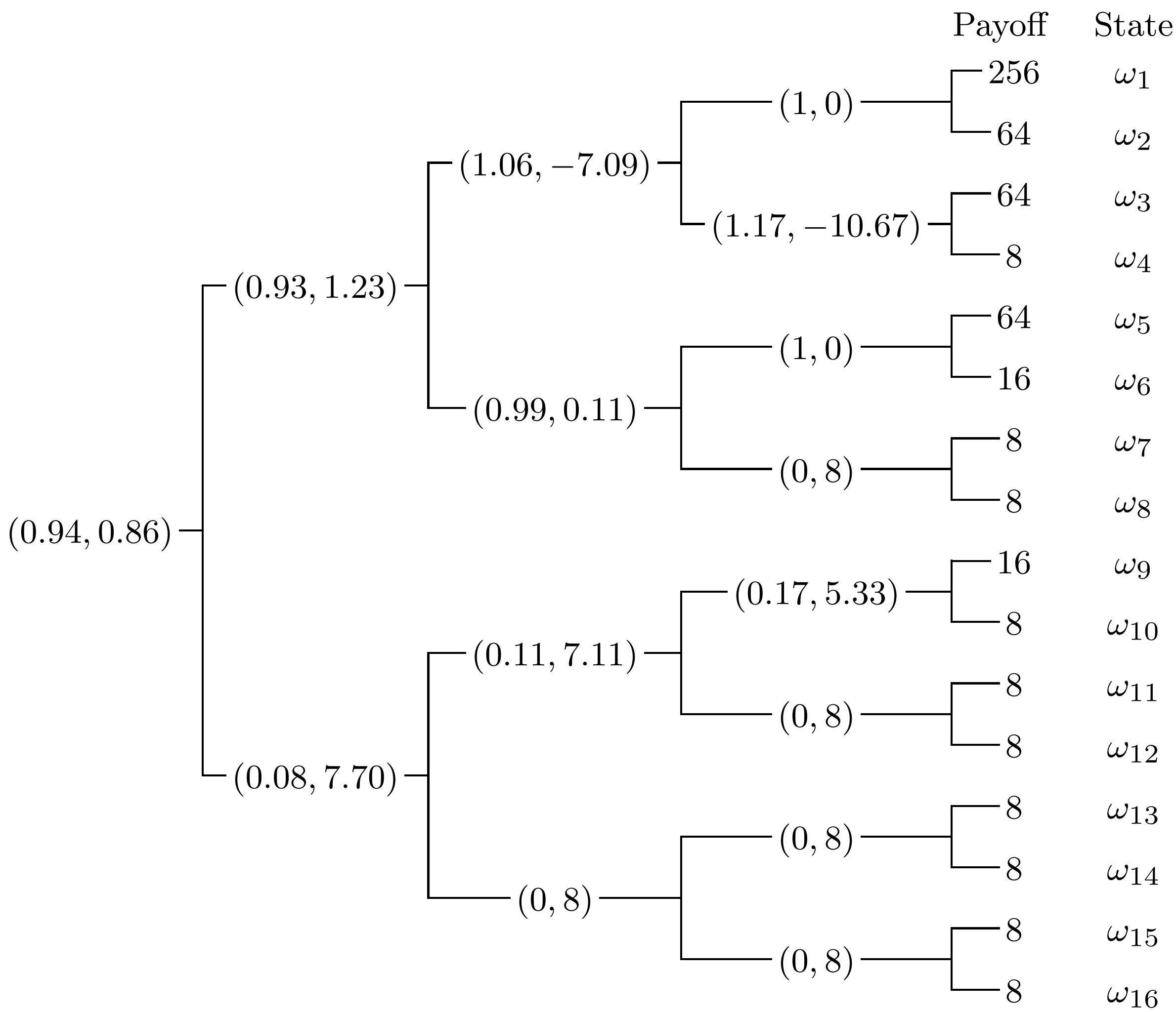
Definition 1. *Let the transaction costs be k. A dynamic self-financing portfolio strategy (Δ, B) is efficient if there is no other dynamic self-financing portfolio strategy (Δ', B') such that:*
(i) *$G(\Delta', B') = G(\Delta, B)$;*
(ii) *$\Delta'_0 S_0 + B'_0 < \Delta_0 S_0 + B_0$.*

In this terminology the stop-loss strategy of Section 2 is inefficient if there are no transaction costs, since the alternative strategy has the same payoff distribution function and has lower initial costs. With transaction costs one can define a stronger form of efficiency similar to the concepts in Edirisinghe, Naik and Uppal (1993) and Bensaid, Lesne, Pagès and Scheinkman (1992) for replicating portfolio strategies. According to Ingersoll (1987) a payoff distribution function $G(\Delta', B')$ stochastically dominates another payoff distribution function $G(\Delta, B)$ if $G(\Delta', B')(x) \le G(\Delta, B)(x)$ for all $x \in \mathbb{R}$ with strict inequality for at least one $x \in \mathbb{R}$.

Definition 2. *Let the transaction costs be k. A dynamic self-financing portfolio strategy (Δ, B) is strongly efficient if it is efficient and there exists no other dynamic self-financing portfolio strategy (Δ', B') such that:*
(i) *$G(\Delta', B')$ stochastically dominates $G(\Delta, B)$;*
(ii) *$\Delta'_0 S_0 + B'_0 \le \Delta_0 S_0 + B_0$.*

Bensaid, Lesne, Pagès and Scheinkman (1992) relax the self-financing condition (1) to

$$\begin{aligned}\big(\Delta_{t-1}(\omega) - \Delta_t(\omega)\big) S_t(\omega) + r B_{t-1}(\omega) - B_t(\omega) \ge \\ k\big|\Delta_{t-1}(\omega) - \Delta_t(\omega)\big| S_t(\omega) \qquad \forall t \le T-1,\ \forall \omega \in \Omega.\end{aligned} \tag{4}$$



Price of strategy: 15.88

$S_0 = 16,\ u = 2,\ d = \frac{1}{2},\ r = 1,\ k = 1\%$

Portfolios: (# of stocks, amount in bonds)

Figure 4: Strongly efficient stop-loss strategy with transaction costs

This means that after each trade some non-negative payoff remains which can be used for consumption if positive. However, when these positive payoffs are invested in riskless bonds until maturity, the final payoff distribution stochastically dominates the final payoff distribution where the positive payoffs are immediately consumed. Hence, if one is checking for strong efficiency and one finds a dynamic portfolio strategy (Δ', B') that satisfies (4) and conditions (i) and (ii) of Definition 2, then there also exists a dynamic self-financing portfolio strategy that satisfies (i) and (ii).

If there are no transaction costs then efficiency implies strong efficiency. Let (Δ, B) be an efficient strategy and assume there exists a strategy (Δ', B') that satisfies (i) and (ii) of Definition 2. Both distribution functions $G(\Delta, B)$ and $G(\Delta', B')$ are stepfunctions, since there are only finitely many possible portfolio values. In Appendix A, it is shown that there exists a dynamic self-financing strategy (Δ'', B'') that has non-negative payoffs in all states and a positive payoff in at least one state such that $G(\Delta, B) = G(\Delta' - \Delta'', B' - B'')$. Since there are no transaction costs $(\Delta' - \Delta'', B' - B'')$ is self-financing. By the no-arbitrage assumption $\Delta_0'' S_0 + B_0'' > 0$ and hence $(\Delta' - \Delta'', B' - B'')$ is more efficient than (Δ, B) in the sense of Definition 1, which gives the required contradiction. Essential in this argument is that the difference between two self financing strategies is again self-financing. This doesn't hold if there are transaction costs.

To check whether a dynamic self-financing portfolio strategy with non-zero transaction costs is efficient there is no easy criterion such as the one described in the previous paragraph for the no transaction costs case. Let (Δ, B) be a dynamic self-financing portfolio-strategy. One might try to find the dynamic self-financing portfolio strategy (Δ', B') among those strategies for which the final payoff distribution dominates $G(\Delta, B)$ and that has the lowest initial costs. This problem can be formulated as a mixed-integer programming problem, as is shown in Appendix B, and its solution will be a strongly efficient dynamic portfolio strategy. However, the complexity of this problem allows one only to find solutions for small values of T. In Figure 4 the optimal strategy in the 4 period case is given where transaction costs are 1% and the strategy must dominate the final payoff distribution of a stop-loss strategy with no transaction costs. Hence, the final payoff should be at least 256 with a probability of 1/16, 64 with a probability of 4/16, 16 with a probability of 6/16, and always exceed 8.

Figure 5 gives the replicating portfolio for the alternative strategy of Figure 3 with 1% transaction costs. Figures 4 and 5 have interchanged payoffs in states ω_4 and ω_6. The strategy given in Figure 5 has the cheapest initial costs of all strategies that generate the payoff random variable of Figure 5. This will be seen more clearly in the next section, where it is shown that there exists a unique "cheapest" strategy that generates this payoff random variable. However, we see that the initial cost of this strategy is higher. In the no transaction costs case both payoff schemes would have the same price, since we interchanged only the payoffs between states with the same state-price density. Hence, with transaction costs we might not only have to change payoffs between states with different state-price densities, but also between states with the same state-price densities. This illustrates the complexity of the optimization problem with transaction costs.

3. Dybvig's Efficient Strategies and Transaction Costs

In this section, we compare the efficient strategies as described in Section 2 with the inefficient stop-loss strategies in the presence of transaction costs. Let (Δ, B) represent a dynamic portfolio strategy that is self-financing if there are no transaction costs. To evaluate how this strategy is affected by transaction costs, we use the following result, based on the proof of Theorem 4 of Boyle and Vorst (1992), for small values of k.

Theorem. *If the transaction costs k on stocks satisfy the following condition*

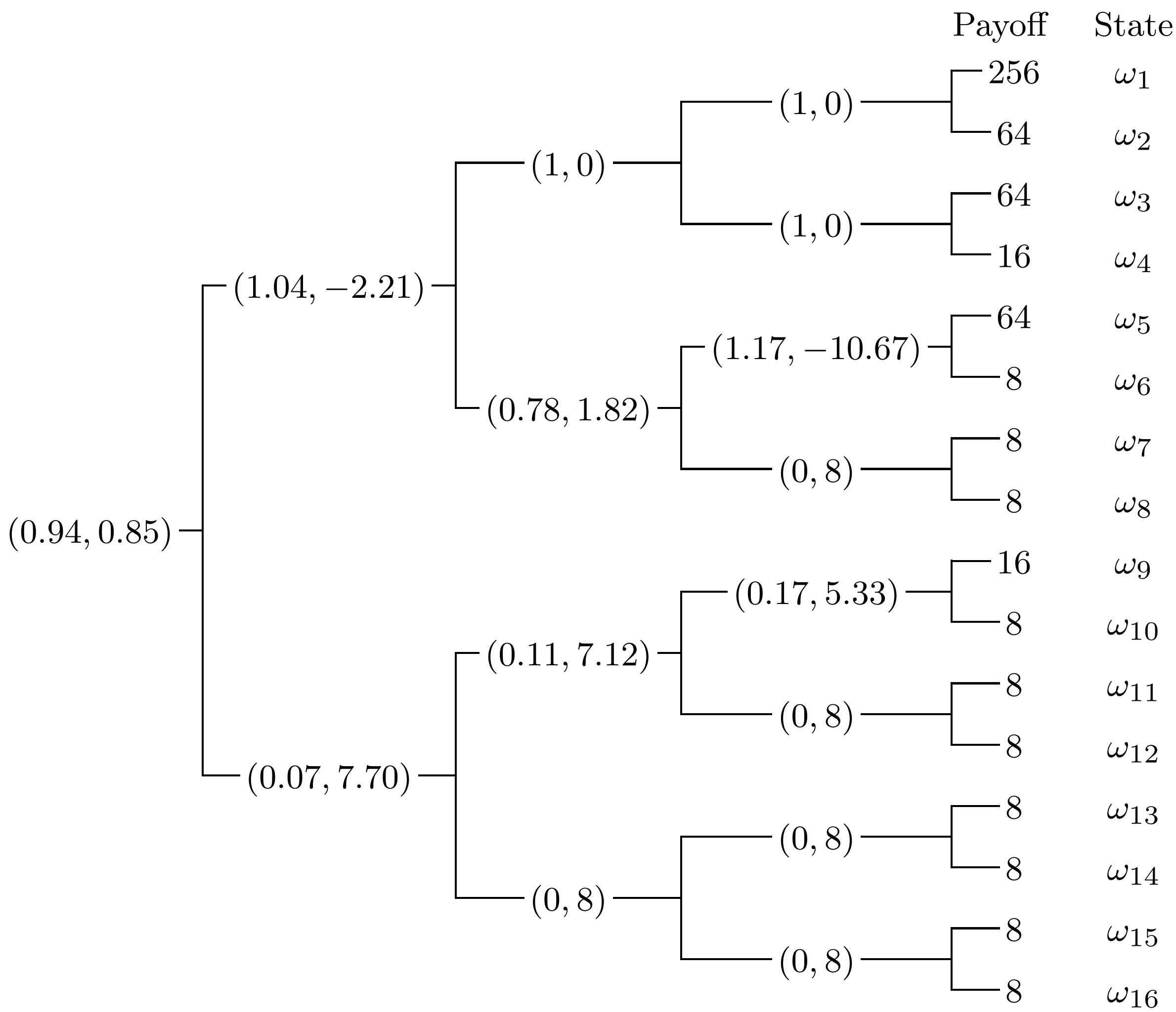
$$(1+k)d < (1-k)u \tag{5}$$

then there exists a unique dynamic portfolio strategy (Δ^k, B^k) that is self-financing for transaction costs k such that

$$\Delta_T(\omega)S_T(\omega) + B_T(\omega) = \Delta_T^k(\omega)S_T(\omega) + B_T^k(\omega) \quad \forall \omega \in \Omega \tag{6}$$

Furthermore, (Δ^k, B^k) can be explicitly found by solving equations (1) *backwards through the expanded binomial tree.*

Proof: In each node of the tree at time t, there are two equations (1), one for the upward-move and one for the downward move. The unknowns

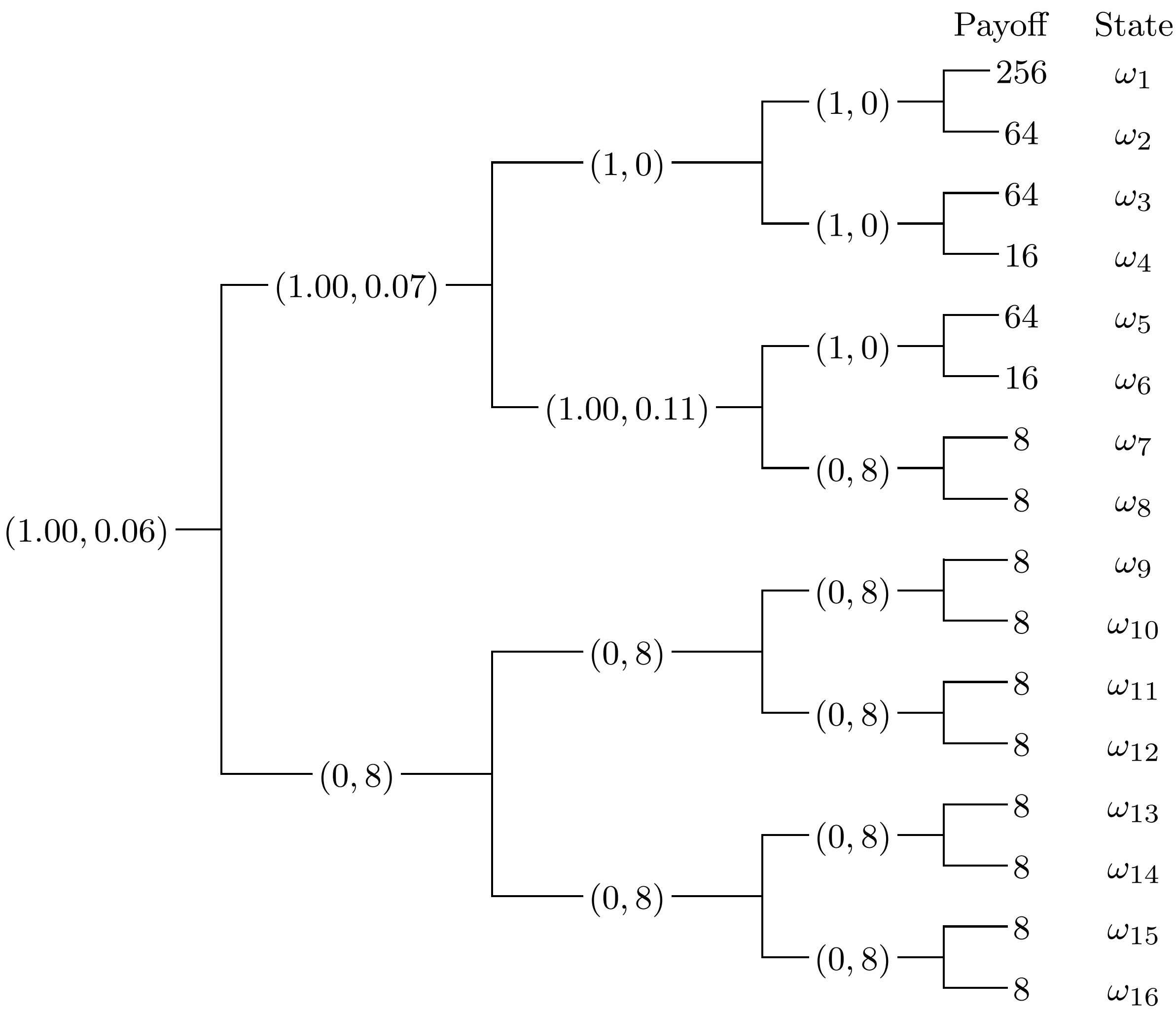


Price of strategy: 15.89

$S_0 = 16,\ u = 2,\ d = \frac{1}{2},\ r = 1,\ k = 1\%$

Portfolios: (# of stocks, amount in bonds)

Figure 5: Alternative stop-loss strategy with transaction costs



Price of strategy: 16.06

$S_0 = 16,\ u = 2,\ d = \frac{1}{2},\ r = 1,\ k = 1\%$

Portfolios: (# of stocks, amount in bonds)

Figure 6: Stop-loss strategy with transaction costs

are $\Delta_t(\omega)$ and $B_t(\omega)$. It follows from the proof of Theorem 4 of Boyle and Vorst (1992) that these two equations with two unknowns have a unique solution if condition (5) is satisfied. These unique solutions can be calculated backwards to find the strategy (Δ^k, B^k). ∎

This theorem allows us to calculate the price of the initial portfolio. This is necessary to be able to perform the strategy under transaction costs in order to obtain the final portfolio (Δ_T, B_T). Figure 6 gives the strategy $(\Delta^{0.01}, B^{0.01})$ for the stop-loss strategy, while in Figure 5 the strategy $(\Delta^{0.01}, B^{0.01})$ is given for the efficient strategy from Figure 3. (Hence the transaction costs are 1%.) We see that the efficient strategy from Figure 3 has higher total transaction costs $(15.89 - 15.80 = 0.09)$ than the stop-loss strategy $(16.06 - 16.00 = 0.06)$.

However, the efficient strategy still has lower initial costs and is thus more efficient than the stop-loss strategy in this simple example. To investigate the influence of transaction costs further, we have to expand the model to more trading intervals, where the changes in stock prices are smaller. Since in each step the number of nodes is doubled in an expanded tree, one can only use a modest number of trading intervals. In Table I we give the initial costs and portfolio compositions both for the stop-loss strategy and the alternative efficient strategy for 18 trading intervals and for several levels of transaction costs. In this example, the values of the volatility, interest rates, time to maturity and expected return on the stock are equal to the values used in Dybvig (1988b). The stop-loss value is 14.4, which is 90% of the original portfolio value. Hence, if the stock price falls below the discounted value of 14.4, the stock is sold for bonds. We see from Table I that not only the total transaction costs differ but also that the so-called efficient dynamic portfolio strategy now requires higher initial total costs than the simple stop-loss strategy, even if the transaction costs on stocks are as low as 0.5%. We conjecture that the effect is more pronounced if we extend to even more trading intervals. Hence, we can conclude that the so-called efficient dynamic portfolio strategy is less efficient than the stop-loss strategy in the presence of transaction costs.

The initial costs of a lock-in strategy are also given in Table I. If an investor follows a lock-in strategy, he initially buys a stock and holds the stock until it reaches a certain pre-specified level (in our case the discounted value of 18.4). At that moment the stock is sold and the revenues are invested in riskless bonds. Hence, a lock-in strategy is

Table I: Prices of replicating strategies using stocks

	Stop-loss at 14.4	
k	Normal	Efficient
0%	16.00 $\begin{cases} 1.0000 & \text{s} \\ 0.0000 & \text{b} \end{cases}$	15.97 $\begin{cases} 0.8849 & \text{s} \\ 1.8119 & \text{b} \end{cases}$
0.5%	16.02 $\begin{cases} 0.9906 & \text{s} \\ 0.1732 & \text{b} \end{cases}$	16.10 $\begin{cases} 0.8726 & \text{s} \\ 2.1339 & \text{b} \end{cases}$
1%	16.06 $\begin{cases} 0.9817 & \text{s} \\ 0.3420 & \text{b} \end{cases}$	16.19 $\begin{cases} 0.8622 & \text{s} \\ 2.3965 & \text{b} \end{cases}$
2%	16.11 $\begin{cases} 0.9656 & \text{s} \\ 0.6591 & \text{b} \end{cases}$	16.53 $\begin{cases} 0.8413 & \text{s} \\ 2.8902 & \text{b} \end{cases}$
	Lock-in at 18.4	
k	Normal	Efficient
0%	16.00 $\begin{cases} 1.0000 & \text{s} \\ 0.0000 & \text{b} \end{cases}$	15.89 $\begin{cases} 0.5796 & \text{s} \\ 6.6185 & \text{b} \end{cases}$
0.5%	16.06 $\begin{cases} 1.0188 & \text{s} \\ -0.2420 & \text{b} \end{cases}$	16.10 $\begin{cases} 0.5534 & \text{s} \\ 7.2484 & \text{b} \end{cases}$
1%	16.12 $\begin{cases} 1.0362 & \text{s} \\ -0.4541 & \text{b} \end{cases}$	16.34 $\begin{cases} 0.5152 & \text{s} \\ 8.0920 & \text{b} \end{cases}$
2%	16.28 $\begin{cases} 1.0737 & \text{s} \\ -0.8994 & \text{b} \end{cases}$	16.92 $\begin{cases} 0.3355 & \text{s} \\ 11.5482 & \text{b} \end{cases}$

$S_0 = 16,\ \sigma = 20\%,\ \mu = 16\%,\ R = 8\%,\ T = 1$

18 trading intervals

similar to a stop-loss strategy, differing only in the pre-specified level. For the lock-in strategy the pre-specified level lies above the initial stock price, while for the stop-loss strategy, it lies below the initial stock-price. Without transaction costs the lock-in strategy is also inefficient and we see from Table I that there exists an efficient strategy that requires only an initial investment of 15.89. This strategy was also constructed using Dybvig's method described in Section 2. In the presence of transaction costs we see that the efficient strategy becomes costly. Therefore it requires substantially higher initial investments than the lock-in strategy and hence is no longer efficient. Thus, the general conclusion is the same for this strategy as for the stop-loss strategy.

4. Options and Cost Reduction in Portfolio Strategies

In the absence of transaction costs the possibility to trade in options does not add extra opportunities in the binomial model. If options are priced arbitrage-free then all portfolio strategies that can be executed with bonds, stocks and options on the stocks, can also be executed with only bonds and stocks, with the same costs. Hence, options are redundant assets. In this section we investigate whether options add extra possibilities or reduce costs in the case where transaction costs are present.

We assume that options are priced at their Cox-Ross-Rubinstein (CRR) binomial tree value, even in the case that there are transaction costs for trading in stocks. Therefore, we do not use the pricing methodology of Boyle and Vorst (1992), which prices long and short option positions using a replicating strategy with transaction costs. We assume that there is an equilibrium between demand and supply in the option markets such that the price of options is the no transaction costs replicating price.[3] This implies that we do not introduce arbitrage opportunities, since the costs of replicating a long option with transaction costs are higher than the market price of this option. For a short option, the costs are lower. However, we assume that there are proportional transaction costs associated with trades in options. Let l be the proportionality factor, which can be different from k, the proportional transaction costs for trades in stocks. Let $C_t(\omega)$ be the price of a specific option on the stock with expiration date T. A dynamic self-financing trading strategy

[3] Alternatively stated: all traders in the market use a CRR-model to establish a fair price for options.

in this option consists of two $\mathcal{F}$-adapted processes Γ and B such that the following equation holds

$$\begin{aligned}\big(\Gamma_{t-1}(\omega) - \Gamma_t(\omega)\big)C_t(\omega) + rB_{t-1}(\omega) - B_t(\omega) &= \\ l\big|\Gamma_{t-1}(\omega) - \Gamma_t(\omega)\big|C_t(\omega) \qquad &\forall t \leq T-1,\ \forall \omega \in \Omega. \quad (7)\end{aligned}$$

$\Gamma_t(\omega)$ is the number of options in the portfolio during the period from t to $t+1$, while $B_t(\omega)$ is the amount invested in riskless bonds during the same period. One can define efficient and strongly efficient strategies for options similar to the concepts for stocks as defined in Section 2. However, to find optimal strategies one would run into the same complexity problems as in the case of stocks. Therefore, our goal in this section is more modest. We want to investigate whether options enable us to replicate the same payoff schemes with lower initial costs. First, we need to investigate whether trading in the option allows us to attain every arbitrary payoff function $V(\omega)$ as is the case for stocks if condition (5) is met. Similar to the proof of the theorem of Section 4, one can show that if

$$(1+l)C_{t+1}(\tilde{\omega}) < (1-l)C_{t+1}(\omega) \tag{8}$$

for all $t \leq T-1$ and all states $\omega, \tilde{\omega}$ that are identical except for the period $t+1$, where $\tilde{\omega}$ is the state in which the stock price path moves downward during this period and ω the the state in which the path moves upward[4], then all payoff functions can be replicated.

For most options, condition (8) will not be satisfied for all $t \leq T-1$ and all states. For example, if the option is far out-of-the-money at some date and at some specific state (for a call option in one of its lower states) it will have a zero value not only at that state, but also at the next period in both possible states. In that case, condition (8) is not fulfilled. Hence, with most options we cannot replicate all payoff schemes. But condition (8) is only a sufficient condition to replicate all payoff schemes. It is therefore certainly possible that a specific payoff scheme can be replicated by a specific option. For example, in the payoff

[4] Using the notation introduced in Section 3 we can write

$$\begin{aligned}\omega &= (\omega_1, \ldots, \omega_t, u, \omega_{t+2}, \ldots, \omega_T)\\ \tilde{\omega} &= (\omega_1, \ldots, \omega_t, d, \omega_{t+2}, \ldots, \omega_T).\end{aligned}$$

scheme of the stop-loss strategy in Table II all payoffs in the lower part of the expanded tree are equal to 8. A call option with exercise price 16 will have value zero at states $\omega_{13}, \omega_{14}, \omega_{15}, \omega_{16}$ at $t = 2, 3$ and 4. However, with this option one can still replicate the stop-loss scheme since in these states the final payments are the same and equations (7) can be solved by not changing the portfolio at these instants. Therefore, we can always find a specific option that enables us to replicate a specific payoff scheme.

In Table II we give the initial costs of the dynamic self-financing portfolio strategies using a call option with an exercise price of 11 to replicate the stop-loss strategy and the alternative efficient strategy. Also, the initial costs of the lock-in strategy and its efficient alternative are given when a put option with an exercise price of 23 is used to replicate the final payoffs. Table II demonstrates that the efficient strategies once again are no longer efficient even if transaction costs are only 0.5%. More importantly, if we compare Tables I and II we see that for all strategies the use of options dramatically decreases the initial costs if the transaction costs on stocks and options are the same, i.e. if $k = l$. Often, the proportional transaction costs on stocks are lower than the proportional transaction costs on options and one would be inclined to use stocks instead of options to replicate a payoff scheme. If we compare row $l = 2\%$ from Table II with row $k = 1\%$ from Table I, we see that the row in Table I has higher initial costs for all four strategies. Thus, even if transaction costs on single trades in options are twice as high as the transaction costs on single trades in stocks, it is still cheaper to use options for the replicating strategies. Hence, options really allow investors to reduce portfolio management costs.

We have only used dynamic self-financing portfolio strategies that use just one option to lower the initial costs and at the same time maintain the payoff distribution. In the real world however, one can use strategies involving not just one stock or one option, but one can replicate with stocks and several put and call options which differ in exercise prices and maturity dates. Hence, the optimal strategy will not only have lower costs than the strategies based on stocks, but also than the strategies based on single options. Therefore, if there are transaction costs, options will enable investors to implement dynamic self-financing portfolio strategies, that reduce initial costs while maintaining the payoff distribution. For large transaction costs this would not be possible if

Table II: Prices of replicating strategies using options

	Stop-loss at 14.4	
l	Normal	Efficient
0%	$16.00 \begin{cases} 1.0064 & \text{c} \\ 10.1110 & \text{b} \end{cases}$	$15.97 \begin{cases} 0.8906 & \text{c} \\ 10.7594 & \text{b} \end{cases}$
0.5%	$16.01 \begin{cases} 1.0039 & \text{c} \\ 10.1309 & \text{b} \end{cases}$	$16.02 \begin{cases} 0.8890 & \text{c} \\ 10.8139 & \text{b} \end{cases}$
1%	$16.01 \begin{cases} 1.0014 & \text{c} \\ 10.1504 & \text{b} \end{cases}$	$16.06 \begin{cases} 0.8874 & \text{c} \\ 10.8629 & \text{b} \end{cases}$
2%	$16.02 \begin{cases} 0.9969 & \text{c} \\ 10.1882 & \text{b} \end{cases}$	$16.12 \begin{cases} 0.8843 & \text{c} \\ 10.9478 & \text{b} \end{cases}$
	Lock-in at 18.4	
l	Normal	Efficient
0%	$16.00 \begin{cases} -1.1035 & \text{p} \\ 21.9204 & \text{b} \end{cases}$	$15.89 \begin{cases} -0.6369 & \text{p} \\ 19.3241 & \text{b} \end{cases}$
0.5%	$16.02 \begin{cases} -1.1048 & \text{p} \\ 21.9501 & \text{b} \end{cases}$	$15.98 \begin{cases} -0.6135 & \text{p} \\ 19.2677 & \text{b} \end{cases}$
1%	$16.05 \begin{cases} -1.1061 & \text{p} \\ 21.9796 & \text{b} \end{cases}$	$16.06 \begin{cases} -0.5847 & \text{p} \\ 19.1972 & \text{b} \end{cases}$
2%	$16.09 \begin{cases} -1.1086 & \text{p} \\ 22.0386 & \text{b} \end{cases}$	$16.25 \begin{cases} -0.5134 & \text{p} \\ 18.9999 & \text{b} \end{cases}$

c : call option with exercise price 11 for stop-loss strategy
p : put option with exercise price 23 for lock-in strategy
$S_0 = 16,\ \sigma = 20\%,\ \mu = 16\%,\ R = 8\%,\ T = 1$
18 trading intervals

only stocks were available to set up the strategy.

5. Conclusion

In this paper we investigated the effect of transaction costs on dynamic portfolio strategies. We showed that even with a modest level of transaction costs of 0.5%, simple trading strategies such as stop-loss and lock-in have a better return distribution than more complex strategies that would be efficient if there were no transaction costs. Furthermore, we showed that the use of options considerably reduces the initial costs of both simple and efficient strategies. This conclusion holds even if the transaction costs on options are twice as high as the transaction costs on stocks. Due to the fact that all calculations require the use of an expanded binomial tree we could only use the binomial model with 18 trading intervals. In further research one should focus on the same results for binomial models with more trading intervals or on similar results for the continuous time model, as was done by Leland (1985), for the replication of options. This will enable us to understand the significance of options for complicated dynamic portfolio strategies better. Especially, a clear view on the cost efficiency of the use of options in an environment with transaction costs is important.

Appendix A

The text of this appendix illustrates that if a dynamic self-financing strategy (Δ', B') stochastically dominates another self-financing strategy (Δ, B) in the no-transaction costs case, then there exists a third dynamic self-financing strategy (Δ'', B'') with non-negative payoffs in all states and a positive payoff in at least one state, such that $G(\Delta, B) = G(\Delta' - \Delta'', B' - B'')$. Let $x_0 = \inf\{y \mid G(\Delta, B)(y) > G(\Delta', B')(y)\}$ and define $\mathrm{Dif} = 2^n(G(\Delta, B)(x_0) - G(\Delta', B')(x_0))$, with n being the number of trading intervals in the binomial tree. Because we assumed that all states have the same probability, each of the 2^n final states has probability 2^{-n}. Therefore, Dif is the number of states that have payoff x_0. Let $x > x_0$ be the next place where there is a jump in the step-function $G(\Delta', B')$ and let ω be one of the states that causes this jump. Hence $\Delta'(\omega)S(\omega) + B'(\omega) = x$. Now reduce the payoff function of (Δ', B') in just this one state with $x - x_0$ to x_0. Let the unique new strategy be denoted by $(\tilde{\Delta}, \tilde{B})$. If $\mathrm{Dif} > 1$, then $x_0 = \inf\{y \mid G(\Delta, B)(y) > G(\Delta', B')(y)\}$, but $2^n(G(\Delta, B)(x_0) - G(\tilde{\Delta}, \tilde{B})(x_0)) = \mathrm{Dif} - 1$ and we can again apply the above procedure and update $(\tilde{\Delta}, \tilde{B})$ once more until $2^n(G(\Delta, B)(x_0) - G(\tilde{\Delta}, \tilde{B})(x_0)) = 0$. Then we find the next $x_0' = \inf\{y \mid G(\Delta, B)(y) > G(\tilde{\Delta}, \tilde{B})(y)\}$. It is easy to check that $x_0' > x_0$ and the x_0's can only be final payoffs of $G(\Delta, B)$ of which there are finitely many. Hence after finitely many steps, we have reduced the payoff of (Δ', B') in several states such that $G(\Delta, B) = G(\tilde{\Delta}, \tilde{B})$. Now let $(\Delta'', B'') = (\Delta' - \tilde{\Delta}, B' - \tilde{B})$ which has all positive payoffs. ■

Appendix B

In this appendix we show that the problem of finding the dynamic self-financing trading strategy (Δ', B') in a world with transaction costs that has the lowest initial costs among the strategies that dominate $G(\Delta, B)$ can be formulated as a mixed-integer programming problem. Let g_i $(i = 1, \dots, m)$ be the different possible payoffs of $G(\Delta, B)$ in the final states and let r_i be the number of final states with payoff g_i. Let $t = 0, \dots, T$ denote the steps in the tree, then we have $j = 1, \dots, 2^t$ states of the world in step t. In the final step $t = T$, we have 2^T final states of the world, hence $\sum_{i=1}^m r_i = 2^T$. If we number the states in step t from top to bottom (as we did in Table I for the final states), then we can reach from state (t, j) the state $(t+1, 2j-1)$ if we go one step up, and state $(t+1, 2j)$ if we go one step down. Finally, the stock price is given by S_{tj} $(t = 0, \dots, T; j = 1, \dots, 2^t)$ in every state.

Decision variables are the real variables

$$\Delta'_{tj}, B'_{tj} \qquad (t = 0, \dots, T; j = 1, \dots, 2^t)$$

and the zero-one variables

$$y_{ij} \qquad (i = 1, \dots, m; j = 1, \dots, 2^T).$$

The optimisation problem can be formulated as follows

$$\min \Delta'_{01} S_{01} + B'_{01} \quad \text{s.t.}$$

$$\left(\Delta'_{tj} - \Delta'_{t+1,2j-1}\right) S_{t+1,2j-1} + rB'_{tj} - B'_{t+1,2j-1} \geq$$

$$k\left|\Delta'_{tj} - \Delta'_{t+1,2j-1}\right| S_{t+1,2j-1} \qquad (t=0,\dots,T-2; j=1,\dots,2^t) \tag{9}$$

$$\left(\Delta'_{tj} - \Delta'_{t+1,2j}\right) S_{t+1,2j} + rB'_{tj} - B'_{t+1,2j} \geq$$

$$k\left|\Delta'_{tj} - \Delta'_{t+1,2j}\right| S_{t+1,2j} \qquad (t=0,\dots,T-2; j=1,\dots,2^t) \tag{10}$$

$$\sum_{i=1}^{m} y_{ij} = 1 \qquad (j=1,\dots,2^T) \tag{11}$$

$$\sum_{j=1}^{2^T} y_{ij} \geq r_i \qquad (i=1,\dots,m) \tag{12}$$

$$\Delta'_{T-1,j} S_{T,2j-1} + rB'_{T-1,j} \geq \sum_{i=1}^{m} g_i y_{i,2j-1} \qquad (j=1,\dots,2^{T-1}) \tag{13}$$

$$\Delta'_{T-1,j} S_{T,2j} + rB'_{T-1,j} \geq \sum_{i=1}^{m} g_i y_{i,2j} \qquad (j=1,\dots,2^{T-1}) \tag{14}$$

$$y_{i,j} \in \{0, 1\} \qquad (i=1,\dots,m; j=1,\dots,2^T) \tag{15}$$

Conditions (9) and (10) are the self-financing conditions in state (t, j) for an up- and a down-move respectively (compare with (4)). (11) assigns each of the final states j to one of the payoffs g_i. (12) guarantees that there are at least r_i states with this payoff. Finally, (13) and (14) require that the final portfolio's $(\Delta'_{T-1,j}, B'_{T-1,j})$ will match the required payoff in the last period T after an up- and a down-move respectively.

11

Summary and Conclusions

DURING the last ten chapters we have been concerned with efficient methods for valuing and managing interest rate and other derivative securities. In the first part of this thesis we have concentrated on interest rate derivatives; in the second part of this thesis we considered other derivative securities like stock options.

In Chapter 2 we gave an introduction to the foundations of option pricing theory in continuous time economies. It is well known that a continuous time economy is free of arbitrage opportunities and is complete if a unique probability measure can be found such that the relative prices in an economy become martingales. Relative prices in an economy are constructed by denominating all prices in terms of a numeraire, which is a marketed asset with strictly positive prices. Usually, we have several marketed assets with strictly positive prices. Therefore the choice of numeraire is, in general, not unique. Both in Chapter 2 and in Chapter 4 it is shown that a different choice of numeraire can greatly simplify the calculation of prices of derivatives.

In Chapter 2 we also introduced another method for valuing derivatives. If an economy is complete, every derivative can be replicated by a trading strategy in the marketed assets. Hence, in the absence of arbitrage opportunities, the value of a derivative can be determined by tracking the value of the replicating portfolio through time. It was shown in Chapter 2 that the value of a replicating portfolio can be described by a partial differential equation. By solving the partial differential equation either analytically or numerically, values for derivatives can be calculated. Several of these solution methods were explored in Chapter 2.

It is clear that the concepts of (no) arbitrage, martingales and partial differential equations are the three cornerstones for the valuation of derivative securities. A good command of all three concepts is needed to get a clear picture of the theory of pricing derivative securities.

Chapters 3 to 7 are concerned with the valuation of interest rate derivatives. In Chapter 3 we showed how the general theory of valuing derivative securities developed in Chapter 2 can be applied to interest rate derivatives. It is not possible to trade in the interest rates themselves. One can only trade in derivative securities like bonds or swaps. Due to the fact that the marketed assets are interest rate derivatives, we can only replicate derivatives with derivatives. This implies that interest rate derivatives cannot be valued "preference-free", but the valuation of interest rate derivatives is still dependent on the market price of risk. However, if interest rate models are fitted to the initial term-structure of interest rates, we can (implicitly) estimate the market price of risk, and the prices of interest rate derivatives are then uniquely determined. It was shown in Chapter 3 that interest rate derivatives can be valued using either equivalent martingale measures, or partial differential equations.

In Chapter 4 we showed that a change of numeraire can provide a powerful tool for explicitly calculating prices of interest rate derivatives. Instead of using the value of the money-market account as a numeraire, it was shown that the value of a T-maturity discount bond is very useful for valuing interest rate derivatives with a payoff at time T. The equivalent martingale measure associated with the T-maturity discount bond is known as the T-forward-risk-adjusted measure. However, using only probabilistic tools it can be difficult to determine explicitly the distribution of interest rates under this new measure. We showed in Chapter 4, that with the help of the partial differential equations, it is possible to explicitly determine the desired distributions.

We also introduced in Chapter 4 the class of normal models. This is a class of interest rate models where the spot interest rate is assumed to be a deterministic function of an underlying (unobserved) process. That is normally distributed. It was shown that several well-known interest rate models fall within this class of normal models. We demonstrated that the only models which are likely to have a rich analytical structure are the models where the distribution of the spot interest rate under the T-forward-risk-adjusted measure is also normally distributed. We proved that this is only the case for normal models where the spot interest rate is either a linear or a quadratic function of the underlying process.

Given the results derived in Chapter 4, we analysed in Chapter 5 the Hull-White model, which is a normal model where the spot interest rate is a linear function of the underlying process. Using the partial differential equation, we explicitly determined the distribution of the spot interest rate under the T-forward-risk-adjusted measure. We proceeded to calculate analytical formulæ for instruments which are widely used in the markets like caps, floors, bond options and swaptions. We also derived an explicit finite difference method to calculate prices for interest rate derivatives for which no analytic formulæ are available.

In Chapter 6 we analysed the squared Gaussian model, which is a model where the spot interest rate is a quadratic function of the underlying process. The squared Gaussian model has the advantage that the spot interest rate never becomes negative. This makes the squared Gaussian model much more realistic than the Hull-White model, where the interest rates can become negative. We proved in Chapter 6 that the squared Gaussian model has an analytic structure as rich as the Hull-White model. Using the techniques of Chapter 4, we determined explicitly the distribution under the T-forward-risk-adjusted measure and showed how prices of interest rate derivatives can be calculated.

Another model which prevents the interest rates from going negative, would be a model where the spot interest rate is the exponential of the underlying process. Since the underlying process is normally distributed, the spot interest rate is log-normally distributed in such a model. However, since the exponential function e^x is neither linear nor quadratic, we know from Chapter 4 that the distribution of the spot interest rate under the T-forward-risk-adjusted measure is not log-normal. In fact, the distribution is unknown, and no analytic formulæ for the prices of interest rate derivatives are available in a log-normal model. Using numerical methods we can still solve the partial differential equation, and obtain prices.

Using historical observations of US-dollar interest rate caps and floors, we empirically compared in Chapter 7 three models: the Hull-White model of Chapter 5, the squared Gaussian model of Chapter 6 and the log-normal model outlined above. All three models are normal models and have a similar underlying process. This underlying process is determined by two parameters. The prices of caps and floors in the three models can be viewed as non-linear functions of these parameters. Hence, we have three non-linear non-nested hypotheses for explaining the prices of caps and floors. These hypotheses can be tested against the data. Due to the fact that the hypotheses are non-linear and non-nested, the standard t-tests are no longer applicable. However, using a special test procedure, known as the P-test, we were able to test the hypotheses against the data. We found that both the Hull-White model and the squared Gaussian model are rejected against the data. The log-normal model was not rejected against the data, and proved to be the superior model. However, we found some evidence suggesting that a model with a distribution of interest rates even more skewed to the right than a log-normal distribution would provide an even better fit.

The second part of this thesis was concerned with efficient methods for valuing and hedging derivative securities like stock or foreign-exchange options.

First, we considered in Chapter 8 efficient methods to calculate hedging parameters from binomial trees. Many people who use binomial trees believe that calculating hedge parameters from nodes within the tree is a "quick and dirty" way of obtaining the hedge parameters delta, gamma and theta. It is generally thought that a finite difference method provides a slow but accurate answer. However, in Chapter 8, we demonstrated that the contrary is true: the calculation within the tree provides the most accurate answer.

In Chapter 9 we showed how a portfolio manager can use options to reshape the return distribution of a portfolio at minimal initial costs. Given an approximation of the return distribution, we proved that the problem at hand is a linear programming problem, which is relatively simple to solve. The methodology used is very general and can be extended in several ways. We demonstrated, for example, that the methodology can be used to find an optimal portfolio in the presence of transaction costs.

In Chapter 10 we relaxed the assumption of the absence of transaction costs, which was made throughout this thesis. We investigated the efficiency of dynamic portfolio strategies in the presence of transaction costs. We showed that dynamic portfolio strategies that are efficient in case no transaction costs are present, but that require frequent rebalancing of the portfolio, are in the presence of transaction costs quickly dominated by very simple (and "inefficient") strategies like stop-loss or lock-in. We also showed how, in the presence of transaction costs, options can be used to enhance the opportunity set of agents in an economy.

References

Arnold,L. (1992). *Stochastic Differential Equations: Theory and Applications.* Reprint Edition, Krieger Publishing Company, Florida.

Beaglehole,D. and M. Tenney (1991). "General Solutions of Some Interest Rate-Contingent Claim Pricing Equations." *Journal of Fixed Income*, Vol. 1, No. 2, pp. 69–83.

Bensaid,B., J.P. Lesne, H. Pagès and J. Scheinkman (1992). "Derivative Asset Pricing with Transaction Costs", *Mathematical Finance*, Vol. 2, No. 2, pp. 63–86.

Black,F. (1976). "The pricing of Commodity Contracts." *Journal of Financial Economics*, Vol. 3, pp. 167–179.

Black,F., E. Derman and W. Toy (1990). "A One-Factor Model of Interest Rates and its Applications to Treasury Bond Options." *Financial Analysts Journal*, January-February 1990, pp. 33–39.

Black,F. and P. Karasinski (1991). "Bond and Option Pricing when Short Rates are Lognormal." *Financial Analysts Journal*, July–August, pp. 52–59.

Black,F. and M. Scholes (1973). "'The Pricing of Options and Corporate Liabilities." *Journal of Political Economy*, Vol. 3, pp. 637–654.

Bookstaber,R. and R. Clarke (1983). "An Algorithm to Calculate the Return Distribution of Portfolios with Option Positions." *Management Science*, Vol. 29, pp. 419–429.

Boyle,P.P. and T. Vorst (1992). "Option Pricing in Discrete Time with Transaction Costs", *Journal of Finance*, Vol. 47, Nr. 1, pp. 271–293.

Brown,S. and P. Dybvig (1986). "The Empirical Implications of the Cox, Ingersoll, Ross Theory of the Term Structure of Interest Rates." *The Journal of Finance*, Vol. 41, pp. 617–630.

Chan,K., A. Karoly, F. Longstaff and A. Sanders (1992). "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate." *The Journal of Finance*, Vol. 47, pp. 1209–1227.

Cox,J., J. Ingersoll and S. Ross (1985). "A Theory of the Term Structure of Interest Rates." *Econometrica*, Vol. 53, No. 2, pp. 385–407.

Cox,J., S.A. Ross and M. Rubinstein (1979). "Option Pricing: A Simplified Approach", *Journal of Financial Economics*, Vol. 7, pp. 229–263.

Cox,J. and M. Rubinstein (1985). *Options Markets*, Prentice-Hall, Englewood Cliffs, New Jersey.

Davidson,R. and J. MacKinnon (1993). *Estimation and Inference in Econometrics.* Oxford University Press, Oxford.

De Munnik,J. (1992). *The Valuation Of Interest Rate Derivatives*, PhD Thesis, Erasmus University Rotterdam.

De Munnik,J. and P. Schotman (1994). "Cross Sectional versus Time Series Estimation of Term Structure Models: Empirical Results for the Dutch Bond Market." *Working Paper*, University of Limburg.

Dothan,M. (1990). *Prices in Financial Markets*, Oxford University Press, New York.

Duffie,D. (1988). *Securities Markets: Stochastic Models*, Academic Press, California.

Duffie,D. (1994). *Transform Methods for Solving Partial Differential Equations*, CRC Press, Boca Raton, Florida.

Dybvig,P. (1988a). "Distributional Analysis of Portfolio Choice", *Journal of Business*, Vol. 61, No. 3, pp. 369–393.

Dybvig,P. (1988b). "Inefficient Dynamic Portfolio Strategies or How to Throw Away a Million Dollars in the Stock Market", *The Review of Financial Studies*, Vol. 1, No. 1, pp. 67–88.

Edirisinghe,C., V. Naik and R. Uppal (1993). "Optimal Replication of Options with Transaction Costs and Trading Restrictions", *Journal of Financial and Quantitative Analysis*, Vol. 28, No. 1, pp. 117–138

Fisher,G. and M. McAleer (1981). "Alternative Procedures and Associated Tests of Significance for Non-Nested Hypotheses." *Journal of Econometrics*, Vol. 16, pp. 103–119.

Garman,M. and S. Kohlhagen (1983). "Foreign Currency Option Values." *Journal of International Money and Finance*, Vol. 2, pp. 231–237.

Geman,H., N. El Karoui and J. Rochet (1995). "Changes of Numéraire, Changes of Probability Measure and Option Pricing." To appear in *Journal of Applied Probability*.

Griffel,D. (1993). *Applied Functional Analysis*, Ellis Horwood Ltd., Chichester.

Harrison,J.M. and D. Kreps (1979). "Martingales and Arbitrage in Multiperiod Securities Markets." *Journal of Economic Theory*, Vol. 20, pp. 381–408.

Harrison,J.M. and S. Pliska (1981). "Martingales and Stochastic Integrals in the Theory of Continuous Trading." *Stochastic Processes and their Applications*, Vol. 11, pp. 215–260.

Heath,D. and R. Jarrow (1987). "'Arbitrage, Continuous Trading and Margin Requirements." *The Journal of Finance*, Vol. 42, pp. 1129–1142.

Heath,D., R. Jarrow and A. Morton (1992). "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation." *Econometrica*, Vol. 60, No. 1, pp. 77–105.

Ho,T., and S. Lee (1986). "Term Structure Movements and Pricing Interest Rate Contingent Claims." *The Journal of Finance*, Vol. 41, pp. 1011–1029.

Hochstadt,H. (1973). *Integral Equations*, John Wiley & Sons, New York.

Hull,J. (1993). *Options, Futures, and Other Derivative Securities* (Second Edition). Prentice-Hall, Englewood Cliffs, New Jersey.

Hull,J. and A. White (1990a). "Pricing Interest-Rate-Derivative Securities." *The Review of Financial Studies*, Vol. 3 (4), pp. 573–592.

Hull,J. and A. White (1990b). "Valuing Derivative Securities Using the Explicit Finite Difference Method." *Journal of Financial and Quantitative Analysis*, Vol. 25 (1), pp. 87–100.

Hull,J. and A. White (1993). "One-Factor Interest-Rate Models and the Valuation of Interest-Rate Derivative Securities." *Journal of Financial and Quantitative Analysis*, Vol. 28 (2), pp. 235–254.

Hull,J. and A. White (1994). "Numerical Procedures for Implementing Term Structure Models I: Single-Factor Models." *The Journal of Derivatives*, Fall 1994, pp. 7–16.

Ingersoll,J. (1987). *Theory of Financial Decision Making*, Rowman & Littlefield, Totowa, New Jersey.

Jamshidian,F. (1989). "An Exact Bond Option Formula." *Journal of Finance*, Vol. 44, pp. 205–209.

Jamshidian,F. (1991). "Bond and Option Evaluation in the Gaussian Interest Rate Model." *Research in Finance*, Vol. 9, 131–170.

Jamshidian,F. (1993). "Bond, Futures and Option Evaluation in the Quadratic Interest Rate Model." *Working Paper*, Fuji International Finance.

Jamshidian,F. (1995). "A Simple Class of Square-Root Interest Rate Models." To appear in *Applied Mathematical Finance*.

Judge,G., C. Hill, W. Griffiths, H. Lütkepohl, T.-C. Lee (1982) *Introduction to the Theory and Practice of Econometrics.* John Wiley & Sons, New York.

Leibowitz,M. and R. Henriksson (1989). "Portfolio Optimization with Shortfall Constraints: A Confidence-Limit Approach to Managing Downside Risk." *Financial Analysts Journal*, Vol. 34 (March–April), pp. 34–41.

Leland,H. (1985). "Option Pricing and Replication with Transactions Costs", *Journal of Finance*, Vol. 40, No. 5, pp. 1283–1301.

Longstaff,F. and E. Schwartz (1992). "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model." *The Journal of Finance*, Vol. 47, pp. 1259–1282.

Lukacs,E. (1970). *Characteristic Functions*, 2nd Edition, Charles Griffin & Co., London.

Merton,R. (1992). *Continuous-Time Finance*, Revised Edition, Blackwell Publishers, Cambridge, Massachusetts.

Mizon,G. and J.-F. Richard (1986). "The Encompassing Principle and its Application to Testing Non-Nested Hypotheses." *Econometrica*, Vol. 54, pp. 657–678.

Pelsser,A. and T. Vorst (1994). "The Binomial Model and the Greeks." *Journal of Derivatives*, Spring 1994, pp. 45–49.

Pelsser,A. and T. Vorst (1995). "Optimal Optioned Portfolios with Confidence Limits on Shortfall Constraints." *Advances in Quantitative Analysis of Finance and Accounting*, Vol. 3, pp. 205–220.

Pelsser,A. and T. Vorst (1995). "Transaction Costs and Efficiency of Portfolio Strategies." To appear in *European Journal of Operational Research.*

Reiner,E. (1992). "Quanto Mechanics." *Risk*, March 1992, pp. 59–63.

Rogers,L. (1993). "Which Model for Term-Structure of Interest Rates Should One Use?" To appear in *Mathematical Finance*, eds. D. Duffie and I. Karatzas, Springer Verlag. [Proceedings of IMA workshop, Minneapolis, June 1993.]

Rogers,L. and D. Williams (1994). *Diffusions, Markov Processes and Martingales* (Second Edition), John Wiley & Sons, Chichester.

Roll,R. (1992). "A Mean/Variance Analysis of Tracking Error", *Journal of Portfolio Management*, Summer 1992, pp. 13–22.

Smith,G.D. (1985). *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, 3rd edition, Oxford University Press, Oxford.

Vasicek,O. (1977). "An Equilibrium Characterisation of the Term Structure." *Journal of Financial Economics*, Vol. 5, pp. 177–188.

Willmott,P., J. Dewynne and S. Howison (1993). *Option Pricing: Mathematical Models and Computation*, Oxford Financial Press, Oxford.

Williams,W. (1980). *Partial Differential Equations*, Clarendon Press Oxford.

Index