Calibration of Interest Rate Term Structure and Derivative Pricing Models

Kin Pang

Submitted for the degree of Doctor of Philosophy

Warwick University

Financial Options Research Centre, Warwick Business School

June 1997

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I am extremely grateful to Stewart Hodges for his constant support, encouragement and advice throughout my work towards this thesis. I would not have been able to produce this thesis without his guidance.

I wish to thank the staff at the Financial Options Research Centre, Warwick Business School. In particular, I would like to express my gratitude to Les Clewlow and Chris Strickland for showing me the ropes. I thank Gail Dean and Asha Leer for their secretarial assistance.

I wish to thank Andrew Carverhill at Hong Kong University of Science and Technology for many fruitful discussions. I thank Nick Webber at Warwick University for introducing me to financial derivatives during my study for my Masters degree at Warwick University.

Finally, I am grateful for funding from the Economics and Social Research Centre between 1993-96 and the Financial Options Research Centre between 1996-97.

ABSTRACT

We argue interest rate derivative pricing models are misspecified so that when they are fitted to historical data they do not produce prices consistently with the market. Interest rate models have to be calibrated to prices to ensure consistency. There are few published works on calibration to derivatives prices and we make this the focus of our thesis.

We show how short rate models can be calibrated to derivatives prices accurately with a second time dependent parameter. We analyse the misspecification of the fitted models and their implications for other models.

We examine the Duffie and Kan Affine Yield Model, a class of short rate models, that appears to allow easier calibration. We show that, in fact, a direct calibration of Duffie and Kan Affine Yield Models is exceedingly difficult. We show the non-negative subclass is equivalent to generalised Cox, Ingersoll and Ross models that facilitate an indirect calibration of non-negative Duffie and Kan Affine Yield Models.

We examine calibration of Heath, Jarrow and Morton models. We show, using some experiments, Heath, Jarrow and Morton models cannot be calibrated quickly to be of practical use unless we restrict to special subclasses. We introduce the Martingale Variance Technique for improving the accuracy of Monte Carlo simulations.

We examine calibration of Gaussian Heath Jarrow and Morton models. We provide a new non-parametric calibration using the Gaussian Random Field Model of Kennedy as an intermediate step. We derive new approximate swaption pricing formulae for the calibration.

We examine how to price resettable caps and floors with the market-Libor model. We derive a new relationship between resettable caplets and floorlets prices. We provide accurate approximations for the prices. We provide practical approximations to price resettable caplets and floorlets directly from quotes on standard caps and floors. We examine how to calibrate the market-Libor model.

1. INTRODUCTION

1.1 BACKGROUND

The investments banks of the 1990's offer to their customers many sophisticated products that were not available in earlier times. Many of these products are financial derivatives, also called options and contingent claims, whose values are derived from the value of fundamental financial variables such as a stock index level or an interest rate. Financial derivatives can be used for many purposes including hedging against adverse movements of specified financial variables and providing an efficient means to express opinions on the movements of financial variables that cannot be so easily achieved without the use of financial derivatives. The importance of financial derivatives to the banking sector is illustrated by their diversity and trading volume. Any bank that can offer these products, and know how to price and hedge their exposures accurately, will have a competitive advantage.

Ever since Black and Scholes (1973) and Merton (1973) laid down the fundamental principles for the pricing of options, banks have been offering increasingly complicated financial derivatives and have been using increasingly sophisticated methods to price financial derivatives. However, despite rapid advancement in our knowledge and ability to model options prices, we will never be able to replicate the real world with all its fine details precisely. Thus options pricing models have to be calibrated to compensate for their deficiencies and to ensure that they comply with market data. We use the term calibration for the process of choosing the parameters that allow the model to agree with market data.

Calibration is central to options pricing in practice. The ease of calibration of a particular options pricing model is as important to practitioners as the model's ability to captured observed behaviour of the relevant financial variables. In the

following chapters we shall focus on the issue of calibration. We restrict our attention to interest rate derivative pricing models although the problems we highlight will apply equally to other areas such as equity and commodity options pricing.

1.2 OBJECTIVE

Banks have been calibrating their options pricing models to market prices and historical data ever since they started applying options pricing models. However, there is little academic work on these issues and the published work is scarce. We wish to contribute to the literature in this area by examining the issues involved, surveying how the different interest rate pricing models can be calibrated and providing suggestions of our own.

1.3 OVERVIEW

The thesis begins in Chapter 2 by defining the general mathematical framework and discussing the modelling issues one has to consider when devising a suitable interest rate derivative pricing model. We distinguish between the two major branches of models, the equilibrium and evolutionary models, and explain why we and practitioners prefer to work with the latter. We also provide a review of the pricing methodology and a review of the various interest rate term structure and derivative pricing models. Finally we review some numerical pricing methods.

Chapter 3 provides a review and analysis of some previous calibration work published in the academic literature.

Chapter 4 examines the short rate models and their calibration. The short rate models are perhaps the simplest of the interest rate models of the term structure. They model how the short rate must evolve to be consistent with an observed interest rate term structure and perhaps selected options prices or a volatility term structure. Short rate models are popular because they can allow

efficient numerical methods and provide analytical tractability. In chapter 4 shows how short rate models can be calibrated easily and highlights some problems inherent to short rate models that may deter their use. We also use short rate models as a basis for analysing model misspecification and its consequences for pricing and hedging.

Chapter 5 examines a model that may have been designed to aid calibration and overcome some of the problems presented by the short rate models. The affine yield model of Duffie and Kan (1996) determines the evolution of an entire interest rate term structure from the joint evolution of a number of reference zero coupon yields. This is certainly attractive, since not only zero coupon yields and their volatilities are relatively easy to ascertain, but also because interest rate derivatives have payoffs defined with reference to only a small number of zero coupon yields. In chapter 5 we show the Duffie and Kan model is unlikely to be used much in practice because it is exceedingly difficult to specify parameters for the model such that the state variables are consistent with being zero coupon yields and even more difficult to calibrate to a volatility term structure or options prices. Furthermore, we illustrate that non-negative Duffie and Kan yield models are in fact equivalent to a class of short rate models we call Generalised Cox, Ingersoll and Ross models that are far easier to calibrate. Anyone wishing to work with a non-negative Duffie and Kan Yield model should start with a Generalised Cox, Ingersoll and Ross model instead.

We will have seen in Chapters 4 and 5 that a major problem with the models considered there is ensuring that the models produce realistic volatility factors for the zero coupon yields. Chapter 6 considers the general approach of Heath, Jarrow and Morton (1992) that models interest rates more realistically than the short rate models. We show how we can price derivatives in the Heath, Jarrow and Morton framework and show how the Heath, Jarrow and Morton approach is

inappropriate, at least in its full generality, because of its intensive computational requirements. Chapter 6 provides timings to show how calibration would be too slow for practical needs even when tricks are used to speed up computation. Chapter 6 reproduces much of Carverhill and Pang (1995). Chapter 6 concludes with a review of some of the subclasses of Heath Jarrow and Morton models that may permit easier calibration, illustrating the compromise between tractability and realism one usually encounters in modelling.

Chapter 7 examines the calibration of Gaussian Heath, Jarrow and Morton models: the class in which interest rates are normally distributed. Although negative interest rates are possible and undesirable, some interest rate derivatives are not affected significantly by the possibility that negative interest rates may occur. A Gaussian Heath Jarrow and Morton model may be acceptable for this situation and these models are relatively easy to calibrate. However, common approaches have not been very satisfactory and in Chapter 7 we propose a new method that we believe is superior. We argue that it may be preferable to calibrate Gaussian Heath Jarrow and Morton models using the Gaussian Random Field Model of Kennedy (1994) as an intermediate step.

Chapter 8 examines the market-Libor model in which forward Libor rates are strictly positive. We provide a review of the market-Libor model and show how it prices interest rate caps and floors consistently with the market convention. We show how the market-Libor model can be used to price resettable caps and floors. We derive a relationship that gives the price of resettable caplets from resettable floorlets and vice versa. We provide two approximations for the prices of resettable caps and floors. One of the approximation has the attraction that it can price resettable caps and floors directly from market quotes for standard caps and floors without having to calibrate the market-Libor model. We discuss how the market-

Libor model can be calibrated for the other approximation when higher accuracy is needed.

Finally Chapter 9 summarises the thesis and provides suggestions for future research.

2. MODELLING ISSUES AND TERM STRUCTURE DERIVATIVE PRICING REVIEW

2.1 *INTRODUCTION*

In this chapter we review the concept of modelling to explain what models aim to achieve in the context of interest rate term structure modelling and derivative pricing. We provide a review of the contingent claims pricing methodology in modern finance and a review of the different classes of interest rate models published in the academic literature. For calibration to be achieved easily, the models must offer either analytical prices or efficient numerical pricing algorithms for common options. It is important to understand when efficient numerical methods would be available so we also review general pricing methods.

2.2 MODELLING ISSUES

It is possible to consider the pricing of securities within a model of the economy, such as in the equilibrium model of Cox, Ingersoll and Ross (1985a, 1985b), and to deduce prices endogenously as a function of the assumed investor and economic properties. However, such models are unlikely to remain tractable if they are to capture fine details of the real world and to date equilibrium models have provided few solutions to practical problems. Furthermore, the equilibrium models fail to reproduce observed market prices. Instead, the partial-equilibrium approach has provided far more useful models. That approach, also known as the no-arbitrage approach, takes as given the behaviour of financial variables, such as security price processes, to determine prices of contingent claims consistent with no-arbitrage.

An arbitrage opportunity is (i) any trade with price zero but has a payoff that is non-negative and positive with positive probability or (ii) any trade with negative

costs but has a non-negative payoff. Agents are assumed to prefer more to less so that the presence of arbitrage opportunities cannot be consistent with an economic equilibrium because all agents will set up arbitrarily large positions in any arbitrage opportunities. Thus the central assumptions, besides the usual perfect market assumptions, are the behaviour of the underlying variables of contingent claims.

The no-arbitrage models transform the user's believes about the underlying financial variables to prices for contingent claims. They allow the user to concentrate on observable quantities rather than the difficult to quantify variables such as risk-premium and expected returns. The numerous contingent claim pricing models differ by how they allow the uncertainty of underlying financial variables to develop. Typically, the underlying variables are assumed to follow certain stochastic processes. Researchers compromise between assumptions that are realistic and those that allow analytical tractability. The fundamental assumption for the pricing of contingent claims is the ability to construct a hedging portfolio, consisting of other securities, that through continuous trading replicates the payoff of the target contingent claim. Thus it is important to be able to capture how the underlying securities or variables evolve through time in a model. The more realistic and similar the assumptions are to the observed behaviour of the real world, the more faith we will have on the model and the prices and hedging ratios it produces. However, making realistic assumptions can make a model intractable so that analytic prices for most common contingent claims cannot be produced forcing the use of numerical methods that are often very slow. In the exacting world of banking, most practitioners and traders need to respond to situations quickly and cannot wait long for numerical routines to produce prices. They are therefore sometimes forced to accept a simple model for its speed in favour of a more complicated and perhaps more realistic model.

Making simplifying assumptions may or may not be appropriate that depends on the eventual use of the model. The chosen model need to be able to represent the key factors determining a contingent claim's value. Other secondary variables can be safely ignored if they do not play a significant part.

For example, consider the pricing of an European option on a pure discount bond. One-factor models of the interest rate term structure are easy to work with but they imply perfect correlation of changes to different interest rates. But, for the pricing of European call and put options on a pure discount bond, the key variable is the variance of the underlying pure discount bond at the option maturity. The imperfect correlation of different interest rates does not play a direct role and so one-factor models would be suitable for the pricing of European call and put options on pure discount bonds. Thus for the pricing of options on pure discount bonds, one-factor models may be preferred because they are easier to calibrate and to work with. They are not, however, suitable for pricing options on coupon bonds where the imperfect correlation of different interest rates is important. Multifactor models of the interest rate term structure can potentially price both options on pure discount bonds and on coupon bonds successfully. However, multifactor models require the estimation of more variables and usually more time to produce prices.

As another example, consider Gaussian interest rate models which allow negative interest rates. It is clear that Gaussian interest rate models, which permit negative interest rates, contravene empirical observations. However, for the pricing of interest rate caps, options that have payoffs when interest rates are high, the possibility that interest rates may become negative is an unimportant property. Gaussian interest rate models offer some analytical tractability and are easy to calibrate. Thus for some applications, Gaussian interest rate models may be preferred to non-negative interest rate models. For the pricing of interest rate

floors, options that have payoffs when interest rates are low, Gaussian interest rate models will be inappropriate if market interest rates cannot be negative. Practitioners therefore need to understand the weaknesses of different models when choosing between the models. They also need to consider the important issue of calibration.

Calibration is the process of finding model parameters to allow the chosen model to best fit market data. This may involve calibrating to ensure that the models are consistent with interest rate dynamics or to ensure that the modes are consistent with a chosen set of market prices. However, calibrating option pricing models to historical data is usually unsatisfactory. Even if parameters could be estimated without error from historical data, historic methods fail for three major reasons. Firstly, historically based methods are always backward looking whereas option prices are based on future events and are forward looking. Secondly, even the more complicated model that currently exist and the new models that will be developed will inevitably fail to capture the full complexity of the real world and so there will always be some degree of misspecification. Thirdly, option prices are determined using equivalent probability measures. Under the alternative probability measures, the underlying variables behave differently from what we postulate for the real world. Therefore we cannot take all parameters from historical estimates and use them with the options pricing models. Under the commonly used risk-neutral measure, however, the volatilities are the same. The historic volatilities should be used as guides for appropriate implied volatilities. Any significant deviations of implied volatilities from historical volatilities would be an indication of severe misspecification and it would be unlikely that the misspecified model would be able to price any options dissimilar to the calibration set. The calibrated model would only serve well as an interpolation tool for options that are similar to the calibration set but fail to price dissimilar accurately because

the calibrated model failed to capture the empirical dynamics. This is particularly important when a model is calibrated to a set of options that will be used to hedge a more complex variety of options.

The calibration problems highlighted above are characteristic of all options pricing models. Thus it will always be necessary to fit implied parameters that typically involves an optimisation to minimise some objective function such as the weighted sum of squared differences between model and market prices. This has strong implications for the choice of models because an optimisation may require many repeated calculations of prices for different model parameters so that even if a model can produce a price in a minute, it may take many hours to calibrate the model. Therefore it is essential for the model to allow either analytical solutions or at least allow quick pricing algorithms for a variety of traded options. The difficulties posed by the calibration of non-negative interest rate models are considerable and for the pricing of some instruments, practitioners may prefer to work with Gaussian interest rate models. Note that in most models, it is not possible to match the prices of all traded options. Indeed, that would be an undesirable property because market prices contain noises and a model with enough flexibility to fit all prices may run the risk of over-fitting. This is manifested usually by unstable model parameters and poor forecasting performance. Differences between market and model prices may be the result of profit opportunities that result from short term deviation of market prices from fair prices.

2.3 PRICING REVIEW

By considering a stochastic intertemporal economy, where uncertainty is represented by a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t: t \in [0,T]\}, P)$ with $\mathcal{F} = \mathcal{F}_T$, Harrison and Kreps (1979) and Harrison and Pliska (1981) show that the existence

of an unique equivalent martingale measure, (EMM), that renders price processes martingales can be used to price contingent claims consistently and its existence guarantees no arbitrages. The uniqueness of the EMM guarantees that the market for the securities is complete, that is, the payoff to any contingent claim can be replicated by a self-financing dynamic strategy that trades continuously in the underlying securities.

Traditionally, prices are measured in terms of a money account defined by

$$\beta(t) = \exp\left[\int_0^t r(u)du\right]$$

which is the time t value of an investment that invests unit cash, \$1, at time 0 in the money markets and continually rolls it over at the short rate, that is, the rate for borrowing or lending for an instantaneous period. We will use the short rate throughout to refer to the instantaneous rate unless otherwise stated. If $F(T)$ is the price of a contingent claim at time T and if there exists an unique EMM, called the risk-neutral measure, for which the process $\{F(s)/\beta(s), \mathcal{F}_s, s \geq 0\}$, where $\mathcal{F}_s$ is the information revealed up to time s, is a martingale, then the time t value of the contingent claim is given by

$$F(t) = \tilde{E}_t\left[\exp\left(-\int_t^T r(u)du\right)F(T)\right] \tag{2.1}$$

where $\tilde{E}_t$ denotes expectation with respect to the risk-neutral measure conditioned on the information known at time t. Under the risk-neutral measure, prices are given by the expected discounted payoffs where the discounting is with respect to the risk free rate and hence the name. Alternatively, El Karoui, Geman and Rochet (1995) show that if we define the Radon-Nikodym derivative by

$$\frac{dQ^N}{d\tilde{Q}} = \frac{N(T)/N(0)}{\beta(T)/\beta(0)}$$

where $N(t)$ is the value of a numéraire at time t then the value of the contingent claim is given by

$$F(t) = N(t)E_t^N\left[\frac{F(T)}{N(T)}\right] \tag{2.2}$$

where E_t^N denotes the expectation, conditioned on the information at time t, with respect to the EMM that renders prices, measured in terms of the numéraire N, martingales. The results of El Karoui, Geman and Rochet (1995) allow us to transform from the risk-neutral measure to any other EMM induced by taking an alternative numéraire to the money account. Thus we see that as far as derivatives pricing is concerned, we only need to examine the processes for the variables underlying the derivative asset under an EMM rather than the objective probability measure. The examples in this review will focus on the traditional risk neutral measure.

Let $P(t, T)$ denote the time t price of a pure discount bond, (PDB), that matures at time T paying \$1 for sure. Then it follows from equation 2.1 that

$$P(t,T) = \tilde{E}_t\left[\exp\left(-\int_t^T r(u)du\right)\right] \tag{2.3}$$

so that the price of any PDB is given once the risk-neutral process for the short rate is specified, that is, the process followed by the short rate under the risk neutral measure. There is considerable flexibility for the choice of the risk-neutral short rate process. Some are chosen for analytical or numerical tractability and others for realism. Our review will concentrate on no-arbitrage models that take market bond prices as model inputs. Contrast this with equilibrium models where bond prices are model outputs. Equilibrium models become no-arbitrage models if we allow their parameters to be time-dependent and solved to match model bond prices with market prices: The processes are made to evolve in such a way that they become consistent with observed prices.

2.4 MODELLING REVIEW

Before we provide a general review of interest rate term structure modelling, we will define some terminology here.

We will use short rate models to refer to those models where the short rate process is Markovian and the number of state variables is the same as the number of factors. This definition of short rate models encompasses all those models that were originally introduced in the literature by specifying the short rate process directly. For example, in one-factor short rate models, the only state variable is the short rate itself. A two-factor short rate model may have the short rate and the level that the short rate reverts to for its state variables. We shall see that the Markovian property of short rate models are particularly attractive from an implementation point of view.

Our definition of short rate models exclude all but very special cases of the models introduced by Heath, Jarrow and Morton (1992), (HJM). Section 2.4.2 reviews the HJM modelling approach. HJM model interest rate dynamics by modelling how the instantaneous forward rate term structure evolves through time. However, as Section 2.4.2 shows, every HJM model can be reformulated to re-express the model as a process for the short rate. Usually the HJM short rate process with be non-Markovian but it does not have to be. The HJM short rate process can be made Markovian by carefully selecting the instantaneous forward rate volatility factors. We shall see in Chapter 6 that there are HJM models that are not short rate models, according to our definition, even though there are state variables because there are more state variables than there are factors.

2.4.1 Short Rate Models

The earliest short rate models were equilibrium models. These include the well known models of Vasicek (1977) and Cox, Ingersoll and Ross (1985b). They do not, however, reproduce the entire interest rate term structure. We shall review

no-arbitrage short rate models of the term structure that are constructed to be consistent with the interest rate term structure. Some no-arbitrage short rate models are obtained by taking the risk-neutral short rate processes from equilibrium models and introducing time dependent parameters to make them consistent with the interest rate term structure. Alternatively no-arbitrage short rate models can be created by specifying the risk neutral short rate process directly provided PDB prices measured in units of the money account are martingales.

2.4.1.1 One-factor No-Arbitrage Short Rate Models

These models typically postulate an Itô process for the short rate where the uncertainty is driven by a Wiener process, $\widetilde{W}(t)$:

$$dr(t) = \mu(r,t)dt + \sigma(r,t)d\widetilde{W}(t)$$

The tilde on the Wiener increment is a reminder that we are considering a process with respect to the risk-neutral measure. $\mu(r, t)$ and $\sigma^2(r, t)$ are functions of the short rate that represent, respectively, the expected instantaneous drift and instantaneous variance rate of the short rate. $\sigma(r, t)$ is known as the absolute volatility of the short rate. The time dependence in the drift and volatility increase arbitrarily the degree of freedom in the short rate process to allow the proposed model to fit specified properties such as the interest rate term structure, and the volatility term structure of interest rates which shows how zero coupon yield volatilities vary with their maturities. There are three major models in this category:

Black, Derman and Toy (1990): $$d\ln r(t) = \left[\theta(t) - \frac{\sigma'(t)}{\sigma(t)}\ln r(t)\right]dt + \sigma(t)d\widetilde{W}(t);$$

Black and Karasinski (1991): $$d\ln r(t) = [\theta(t) - \phi(t)\ln r(t)]dt + \sigma(t)d\widetilde{W}(t);$$

Hull and White (1990, 1993): $dr(t) = \left[\theta(t) - \phi(t)r(t)\right]dt + \sigma(t)r(t)^{\beta} d\widetilde{W}(t)$.

We examine these models in more detail in Chapter 4. Hogan and Weintraub (1993) show that the Black, Derman and Toy (1990), (BDT), Black and Karasinski (1991), (BK) and Hull and White (1990, 1993) model with $\beta = 1$, are unsatisfactory because they attach negative infinite values to Eurodollar future contracts. One factor interest rate models are also inappropriate for valuing contingent claims that are sensitive to changes to slopes of the zero coupon yield term structure because the one-factor assumption implies all interest rate changes are perfectly correlated.

2.4.1.2 Multifactor No-Arbitrage Short Rate Models

Multifactor models are needed to price and hedge derivatives that are sensitive to the correlation structure of interest rate changes. Practitioners have a wide variety of multifactor models available to them that includes the multifactor extensions of the models of the one factor short rate models. Others include Langetieg (1980), Fong and Vasicek (1991), Longstaff and Schwartz (1992) and Chen (1994) when the parameters are allowed to be time dependent..

Most multifactor models belong to the class of models considered by Duffie and Kan (1996). The only notable exceptions are the lognormal models of BDT and BK. Duffie and Kan (1996) consider models that are characterised by

$$r = f + \underline{G}^T \underline{X}$$

where $\underline{X}$, the state variables, solve

$$d\underline{X} = (a\underline{X} + \underline{b})dt + \Sigma \begin{bmatrix} \sqrt{v_1(\underline{X})} & & 0 \\ & \ddots & \\ 0 & & \sqrt{v_n(\underline{X})} \end{bmatrix} d\underline{\widetilde{W}}$$

where $v_i = \alpha_i + \underline{\beta}_i^T \underline{X}$, $i = 1$ to n.

Duffie and Kan (1996) show the zero coupon yields in their model are an affine function of the state variables. We examine Duffie and Kan Models in detail in Chapter 5.

2.4.2 HJM Approach to Term Structure Modelling

Bond prices can be expressed as

$$P(t,T) = \exp\left[-\int_t^T f(t,u)du\right]$$

where $f(t, u)$ is the instantaneous forward rate one can contract at time t to borrow or lend at time u for an instant. HJM (1992) assume a family of forward rate processes that under the risk-neutral measure is given by

$$f(t,T) = f(0,T) + \int_0^t \alpha(v,T,\omega)dv + \int_0^t \underline{\sigma}(v,T,\omega)\cdot d\underline{\widetilde{W}}(v)$$

where $d\underline{\widetilde{W}}$ is a vector of independent Wiener increments in the risk-neutral measure, $\underline{\sigma}(v, T, \omega)$ are the volatility factors and $\alpha(v, T, \omega)$ is the drift. Both $\underline{\sigma}(v, T, \omega)$ are $\alpha(v, T, \omega)$ may depend on the path of the Brownian motion up to time t. HJM (1992) show the drift is constrained by no-arbitrage to be

$$\alpha(s,u,\varpi) = \underline{\sigma}(s,u,\varpi)\cdot\int_s^u \underline{\sigma}(s,v,\varpi)dv\,.$$

The short rate $r(t)$ is given by

$$r(t) = \lim_{T\to t} f(t,T)$$

and it follows that the risk-neutral process for the short rate, suppressing the dependence of the forward rate volatilities on the Brownian path, is given by

$$dr(t) = \left\{\frac{\partial f(0,t)}{\partial t} + \int_0^t\left[\frac{\underline{\sigma}(u,t)}{\partial t}\cdot\int_u^t \underline{\sigma}(u,v)dv + |\underline{\sigma}(u,t)|^2\right]du + \left[\int_0^t \frac{\underline{\sigma}(u,t)}{\partial t}\cdot d\underline{\widetilde{W}}(u)\right]\right\}dt + \underline{\sigma}(t,t)\cdot d\underline{\widetilde{W}}(t)$$

which is in general non-Markovian even when the volatility factors are deterministic because the third term of the drift depends on the Brownian path up to time t.

Notice that once the volatilities are specified the risk-neutral process is completely specified. Furthermore, the volatilities are the same under both the risk-neutral and objective measures. Thus the HJM approach appears to be particularly attractive and the observation suggests a simple calibration procedure based on historical data. However, historically based methods will fail in practice to produce options prices consistently with market quotes due to the reasons we have discussed in Section 2.1. We examine a popular historic method and introduce a preferred implied method in Chapter 7.

2.5 PRICING METHODS

There are basically two approaches to pricing derivatives: The first approach, which we call the martingale pricing approach, corresponds to solving equation 2.2 which shows that if contingent claims has value $F(T)$ at time T, with no intermediate payoffs, then the current value is given by the expected discounted value under the risk-neutral measure. If there are intermediate payoffs, then those payoffs can be rolled over to time T by investing the intermediate payoffs in T maturity PDB so that taking the expected discounted value under the risk-neutral measure still applies.

The second approach, which we call partial differential equation pricing approach, can be applied if there are state variables in the model for if they exist, then it may be possible to express the value of the derivative as the solution to a partial differential equation.

Very often it is not possible to obtain closed form solutions for either approaches and we have to resort to numerical methods. After the existence or

otherwise of closed form solutions for prices of common derivatives, the existence or otherwise of efficient numerical methods for the different interest rate models play a primary role in determining their practicality.

When we cannot find analytical solutions to equation 2.2 in the martingale pricing approach, we can evaluate the expectation numerically using trees or Monte Carlo simulations. Chapter 4 considers short rate trees in detail. Basically, we can construct a tree to generate approximate discrete distributions of the underlying variables through time according to their processes under the chosen EMM. We can propagate a tree to the option maturity to give a discrete approximation to the terminal option payoff distribution. Then, in the case of the risk neutral measure, we can evaluate the expected discounted option payoff within the tree to provide an estimate of the option value. Chapter 6 considers Monte Carlo simulations in detail. In basic Monte Carlo methods we simulate risk-neutral paths for the underlying variables to the option maturity and evaluate the discounted option payoffs. Each simulation provides a random sample from the discounted option payoff distribution. We generate many paths and take the average of the discounted option payoffs to give an unbiased estimate of the option value.

In the partial differential equation pricing approach, we can use finite difference methods to solve numerically the partial differential equation for option values when we cannot solve the partial differential equation analytically. There is a large literature on numerical solutions of partial differential equations. For our brief review of pricing, it is sufficient to say that basically, finite difference methods systematically assign option values across a time and states space grid such that the finite difference approximation to the partial differential equation and boundary conditions are satisfied at all grid nodes. The partial differential approach returns the assigned value to the node corresponding to the current time

and states for the option value. See Smith (1975) for more details on how to solve partial differential equations numerically and Clewlow (1992) for a review of finite difference methods applied to option valuation problems.

To price American options, both the pricing approaches as described have to be modified. In the martingale pricing approach, equation 2.2 has to be modified to

$$F(t) = \sup_{\tau \in \Psi[t,T]} E_t^N \left[\frac{N(t)}{N(\tau)} F(\tau) \right] \tag{2.4}$$

where $\psi[t, T]$ is the class of all early exercise strategies, that is, the value of American options is maximised over all early exercise strategies. We can still use trees readily for the martingale pricing approach. At each tree node, the value of the American option is assigned the greater of the early exercise value and the expected discounted value over the next time step. This is the only modification needed. Monte Carlo simulation methods cannot be used so easily because it is difficult to determine the early exercise boundary easily but researchers have recently developed Monte Carlo simulation methods that can price American options. For example, see Broadie and Glasserman (1994). They are, however, generally slow and would be inappropriate for calibrating models.

To price American options in the partial differential equation pricing approach, we need to add an early exercise boundary to the partial differential equation: The value of the American option at all the time-space nodes must exceed the early exercise value of the option.

Note that for both approaches to be practical it is important that the models provide state variables. The partial differential pricing approach cannot even be used unless there are state variables so we limit our discussion on efficiency to the martingale pricing approach.

The efficiency of the methods we can use in the martingale pricing approach depend critically on whether there are state variables. This is because for good accuracy it is important that there are many time steps between the current time and the option maturity. This is particularly so for American options because it is important to provide many early exercise opportunities. The number of early exercise opportunities is equal to the number of time steps to the option maturity minus one.

When there are state variables, it may be possible to generate recombining trees to approximate the distributions for the state variables. Trees are recombining if there are more than one way to reach other tree nodes from the initial node, except perhaps to those at the boundaries of the discretised distribution. With non-recombining trees, there is only one path to each node. When we can generate recombining tree the number of tree nodes for a given number of time steps will be far smaller than the case where it is not possible to construct a recombining tree. This is important because we need a large number of time steps and with non-recombining trees, there may be too many tree nodes and too much computation, for the method to be practical.

It is also far easier to conduct Monte Carlo simulations when there are state variables. If there are state variables, then we only need the current values to sample their values a time step further ahead. If a model does not provide state variables, then it would be necessary to examine the entire path taken by the variables to reach their current levels to sample future values.

Our review on pricing methods shows that whether the different interest rate models admit state variables is practically very important. Our review of HJM models noted that their short rate processes are in general non-Markovian. Thus, our review on pricing methods suggest that HJM models may not be able to price complex options efficiently. We examine HJM pricing methods in detail in Chapter

6. If practitioners have to choose between short rate models and HJM models, they would have to consider what advantages HJM models offer that compensate for their lack of efficient numerical methods.

2.6 REFERENCE

1) Black F. and Karasinski P., "*Bond and Option Pricing when Short Rates are Lognormal*", Financial Analysts Journal, July-Aug 1991, pp 52-59.

2) Black F., Derman E . and Toy W., "*A One-Factor Model of Interest Rates and It's Application to Treasury Bond Options*", Financial Analysts Journal, Jan-Feb 1990, pp 33-39

3) Broadie M. and Glasserman P., "*Pricing American-Style Securities Using Simulation*", Working Paper, 1994, Columbia University.

4) Chen L., "*Stochastic Mean and Stochastic Volatility: A Three-Factor Model of the Term Structure and its Application in Pricing in Interest Rate Derivatives*", Working Paper, Federal Reserve Board, October 1994.

5) Clewlow L.J., "*Finite Difference Techniques for One and Two Dimension Option Valuation Problems*", Nov 1992, FORC pre-print 90/10, University of Warwick.

6) Cox J.C., Ingersoll J.E. and Ross S.A., "*An Intertemporal General Equilibrium of Asset Prices*", Econometrica, Vol. 53, No.2, 1995a.

7) Cox J.C., Ingersoll J.E. and Ross S.A.,"A Theory of the Term Structure of Interest Rates", Econometrica, Vol. 53, No.2, 1995b.

8) Duffie D., Kan R., "*A Yield-Factor Model of Interest Rates.*", Mathematical Finance, 1996, Vol. 6, No. 4, pp 379-406.

9) Fong H.G., Vasicek O.A., "*Interest rate Volatility as a Stochastic Factor.*", Working Paper, Gifford Fong Associates, February 1991.

10) Geman H., El Karoui N., and Rochet J.-C., "*Changes of Numéraire, Changes of Probability Measure and Option Pricing*", Journal of Applied Probability, Vol. 32, 1995.

11) Harrison J.M. and Kreps D.M., 1979, "Martingales and Arbitrage in Multiperiod Securities Markets", *J. Econ. Theory* 20, 381-408.

12) Harrison J.M. and Pliska S.R., 1981, "Martingales and Stochastic Integrals in the Theory of Continuous Trading", *Stoch. Proc. Appl.* 11, 215-260.

13) Heath D., Jarrow R., and Morton A.., "*Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation*", Econometrica, Vol. 60, No. 1, 1992.

14) Hull J. and White A., "*One Factor Interest Rate Models and the Valuation of Interest Rate Derivative Securities*", *Journal of Financial and Quantitative Analysis*, Vol. 28 (2), June 1993, pp 235-254.

15) Hull J. and White A., "*Pricing Interest Rate Derivative Securities*", The Review of Financial Studies, Vol. 3 (4), 1990.

16) Langetieg T.C., "*A Multivariate Model of the Term Structure.*", Journal of Finance, March 1980, pp. 71-97.

17) Longstaff F.A., Schwartz E.D., "*Interest-Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model.*", Journal of Finance, September 1992, pp. 1259-1282.

18) Smith G.D., "*Numerical Solution of Partial Differential Equations*", Oxford Mathematical Books, Oxford University Press, 1975.

19) Vasicek O.A., "*An Equilibrium Characterisation of the Term Structure.*", Journal of Financial Economics, Vol. 5 1977, pp. 177-188.

3. LITERATURE REVIEW ON CALIBRATION OF INTEREST RATE TERM STRUCTURE MODELS TO OPTIONS PRICES

3.1 INTRODUCTION

There are many papers, for example Chan et al (1992), in the academic literature that examine parameter estimation for a variety of interest rate term structure models and others, for example Gibbons and Ramaswamy (1993), that test whether the models are consistent with empirical interest rate dynamics. There are, however, relatively few published papers, that we are aware of, that examines the calibration of interest rate derivative pricing models to options prices. This we have already argued is particularly important to practitioners and is the focus of this thesis. In this chapter we review the following three papers on the calibration of short rate models; Hull and White (1993a, 1995 and 1996). We also review the following three papers on the calibration of HJM type models; Amin and Morton (1994), Brace and Musiela (1994) and Brace, Gatarek and Musiela (1995).

3.2 CALIBRATING SHORT RATE MODELS TO OPTIONS PRICES

Although short rate models and short rate tree construction techniques have been published in the academic journals since Black Derman and Toy (1990), very little has been published on their calibration to options prices. Black and Karasinski (1991) say that with three time dependent functions in their model, one of them can be used to fit to cap prices. Little more is said in Black Karasinski (1991) on calibration to options prices. More recently the papers by Hull and White (1993a, 1995, 1996) discuss how their models can be calibrated to options prices.

3.2.1 Hull and White (1993a, 1995, 1996)

The majority of the models considered in these three papers can be summarised by the process

$$dx = [\theta(t) - a(t)x]dt + \sigma(t)dz$$

where $x = f(r)$ is some function of the short rate r. The models proposed by Black, Derman and Toy (1990) and Black and Karasinski (1991) are special cases of the above model. Essentially, the time-dependent functions provide extra parameters to allow the model to be calibrated to more market data. Typically, $\theta(t)$ is used to match the initial interest rate term structure, $\sigma(t)$ is used to determine future volatility of the short rate and $a(t)$ is used to fit the initial volatility term structure of zero coupon yields. Alternatively, $a(t)$ and (or) $\sigma(t)$ can be used to fit options prices. We show how, for a fixed $\sigma(t) \equiv \sigma$, we can calibrate $a(t)$ to caplet[1] prices in Chapter 4. However, when any of $a(t)$ and $\sigma(t)$ is allowed to be time dependent, the volatility term structure of zero coupon yields in general evolves deterministically through time in an unrealistic way. This is a particular problem for the pricing of long maturity options and for options that are sensitive to the shape of the volatility term structure. Thus Hull and White (1995, 1996) recommend that only $\theta(t)$ be time-dependent and that option prices should be fitted by minimising an error function of the differences between model and market prices over $a(t) \equiv a$ and

---

[1] Caplets are instruments that can be used to limit the interest rate charged on a floating loan. For example a caplet may have at time $T+\delta$, the payoff $\delta\,[L(T) - K]^+$ where K is the cap rate and $L(T)$ is the δ-Libor rate, defined by $1 + \delta L(T) = 1/P(T, T+\delta)$, charged on the loan over the period $[T, T+\delta)$. If the realised δ Libor rate is greater than the cap rate, the borrower is compensated by an amount that in effect capped the borrowing rate to K. A cap consists of a portfolio of caplets that cover adjacent periods to limit the borrowing rate to the length of the cap.

$\sigma(t) \equiv \sigma$ subject to some appropriate constraints to ensure the parameters have plausible values. We consider this issue in detail in Chapter 4.

The Hull and White models offer a relatively simple approach to interest term structure modelling and derivatives pricing. It is possible to construct a recombining tree for the short rate easily and price European and American options in a similar fashion to the Cox, Ross and Rubinstein (1979) binomial tree. Furthermore, it is also possible to price some path dependent options with the tree by extending the information stored in the tree nodes. See for example, Hull and White (1993b). The Hull and White models allow practitioners to price some derivatives, such as the path-dependent variety, much more quickly than HJM type models which would in general would require time consuming Monte Carlo simulations.

The Hull and White models do have their disadvantages. Without allowing for time dependence of $a(t)$ or $\sigma(t)$, the ability of the models to fit a large number of options prices is limited. We discuss applications where it is essential that the model is able to price calibration options very accurately in Chapter 4 so that both $\theta(t)$ and $a(t)$ will have to be time-dependent. It is also important that the model can match empirical interest rate dynamics accurately. The single factor Hull and White models fail in this respect because they imply that all zero coupon yield changes are perfectly correlated. Empirical evidence indicate that this is clearly not the case.

There are various multifactor extensions of the models considered here that allow zero coupon yield changes to be imperfectly correlated. In these models, one of the parameters is made time-dependent parameter to ensure the model is consistent with the initial interest rate term structure. The remaining fixed parameters can be varied to provide a wider range of shapes for the volatility term structure and so enable a better fit to options prices. Indeed, most models with

sufficient degrees of freedom can be made to fit a cross section of prices but they may be unreliable for applications other than serving as interpolation tools for options that are similar to the calibration set. The calibrated models should also model interest rate dynamics accurately. It may be difficult to fit the models to options prices without implying unlikely values for other parameters. Increasing the number of factors also increases the computational requirement exponentially so that the computational advantages of the Hull and White models over the HJM approach decreases rapidly. These issues are shared by all short rate models.

3.3 CALIBRATING HEATH, JARROW AND MORTON TYPE MODELS TO OPTIONS PRICES

In contrast to the problems we have just discussed regarding the implied volatility structure in the short rate models, the HJM approach provide much greater flexibility on their specification, except that, of course, they satisfy the conditions given by assumptions C1-C6 in HJM (1992). The volatility factors can be made time stationary. We can also fit a wide range of options prices. The downside is that the HJM approach does not provide option values so easily.

3.3.1 Amin and Morton (1994)

This paper takes an approach that closely resembles the HJM methodology in its original form. Options prices are determined by the evolution of an instantaneous forward interest rate curve that is described by an Itô process rather than the evolution of bond prices that some of the more recent papers adopt. Amin and Morton (1994) assume an one-factor model of the instantaneous forward rates that follow the process

$$df(t,T) = \alpha(t,T,\cdot)dt + \sigma(t,T,f(t,T))d\widetilde{W}(t)$$

with

$$\alpha(t,T,\cdot) = \sigma(t,T,f(t,T))\int_t^T \sigma(t,T,f(t,u))du.$$

Amin and Morton (1994) also assume that the forward rate volatilities take one of the following forms

1. Absolute: $\sigma(\cdot) = \sigma_0$
2. Square Root: $\sigma(\cdot) = \sigma_0 f(t, T)^{1/2}$
3. Proportional: $\sigma(\cdot) = \sigma_0 f(t, T)$
4. Linear Absolute: $\sigma(\cdot) = [\sigma_0 + \sigma_1(T-t)]$
5. Exponential: $\sigma(\cdot) = \sigma_0 \exp[-\lambda(T-t)]$
6. Linear Proportional: $\sigma(\cdot) = [\sigma_0 + \sigma_1(T-t)]f(t, T)$.

These assumptions allow Amin and Morton (1994) to construct a non-recombining forward rate tree as outlined in Heath, Jarrow and Morton (1991), to value interest rate derivative securities. Note that each tree node has to contain sufficient information to produce an entire forward rate curve and since the tree is non-recombining, valuation using the tree requires vast amount of computing resources and a long time to compute. The tree cannot as a result have many time steps. Perhaps for this reason, Amin and Morton (1994) only price Eurodollar futures options with up to two years in maturity. We examine the Heath, Jarrow and Morton (1991) forward rate tree in detail in Chapter 6.

To see how the futures options are priced, let the futures price at date t for a contract that matures at date T be $F_T(t)$. The Eurodollar futures price at maturity T, $F_T(T)$, is given by

$$F_T(T) = 100[1 - L(T)]$$

where $L(T)$ is the reference Libor rate at time T given by the time T forward rate curve. Furthermore $F_T(t)$ is given by

$$F_T(t) = \tilde{E}_t\left[F_T(T)\right].$$

Suppose the option strike is X, then once the forward rate tree has been constructed $F_T(t)$, the maturity payoff $[F_T(T)-X]^+$ and early exercise values $[F_T(t)-X]^+$ follow easily. Thus the American futures options can be priced using backwards induction through the forward rate tree in a similar fashion to Cox, Ross and Rubinstein (1979).

Amin and Morton (1994) minimise the squared errors between market and model prices to extract optimal volatility parameters from the option prices to provide an implied volatility function for each of the six functional forms they consider.

In addition to introducing the concept of implied HJM volatility functions, Amin and Morton (1994) also provide a method to handle what is now called the convexity adjustment for the extraction of forward rates from futures prices. Amin and Morton (1994) extract their initial interest rate term structure from Eurodollar futures prices. It is not possible to extract the forward rates from the time t futures prices without first specifying the volatility structure and it is not possible to price the options without an initial forward rate curve that gives an implied volatility curve. Therefore Amin and Morton (1994) use an iterative two stage scheme that extracts the initial forward rate curve and its volatility term structure. Stage one extracts the forward rate curve for a given volatility structure and stage two extracts a volatility structure for a given forward rate curve. To start of the procedure, Amin and Morton (1994) make an initial assumption for the volatilities that allows the forward rates to be extracted from the Eurodollar futures prices using stage 1. They then feed the extracted forward rate curve into stage 2 to update the volatility parameters by fitting to the Eurodollar futures options prices. The iterative scheme then commences by feeding the volatilities from stage 2 into stage 1 and repeating the cycle until the volatility parameters become stable.

The contributions by Amin and Morton (1994) to the literature include introducing the concept of extracting HJM volatility structures from options prices and highlighting the difficulty of extracting interest rates from Eurodollar futures prices. Their method is however lacking in several areas:

1. Their calibration is slow and unsuitable for practitioners because it uses a non-recombining forward rate tree;
2. Their choice of calibration options with maturities of less than 2 years will not allow their calibrated models to price longer maturity options accurately;
3. Their choice of calibration options does not allow for accurate calibration of multifactor models because Eurodollar futures options prices are insensitive to the correlation structure of zero coupon yield changes.
4. Their technique cannot be applied to multifactor model because it will be much too slow.
5. It is not clear whether the functional forms for their volatilities are appropriate.

3.3.2 Brace and Musiela (1994)

This paper examines calibration of Gaussian HJM to options prices. Using their notation, the time t instantaneous forward rate with maturity x, $r(t, x)$, which is equivalent to $f(t, t+x)$ in the HJM(1992) notation, follows the process

$$dr(t,x) = \left(\frac{\partial}{\partial x} r(t,x) + a(t,x)\right)dt + \underline{\tau}(t,x) \cdot d\underline{\widetilde{W}}(t)$$

where

$$a(t,x) = \underline{\tau}(t,x) \cdot \int_0^x \underline{\tau}(t,u)du$$

and $\underline{\tau}(t, x)$ are the volatility factors for the forward rate $r(t, x)$. Brace and Musiela (1994) calibrate to options on futures on zero-coupon bonds and caplets prices in the Australia market. Essentially, the options on the futures are used to obtain a

good estimate for the volatility structure at the short end. The caplets allow the volatility structure to be extended.

Brace and Musiela (1994) show that if $cpl(t)$ is the time t value of a caplet maturing at time T and paying $\delta\,[L(T) - K]^+$ at time $T+\delta$, where K is the cap rate and $L(T)$ is the spot δ-Libor rate at time T then

$$cpl(t) = \delta\left\{P(t,T)N[-h] - (1+k\delta)P(t,T+\delta)N[-h-\zeta]\right\}$$

where

$$\zeta^2 = Var\left[\log P(T,T+\delta)\,\Im_t\right] = \int_0^{T-t}\left|\int_s^{s+\delta}\underline{\tau}(T-s,u)du\right|^2 ds$$

and

$$h = \frac{1}{\zeta}\left(\log\frac{(1+k\delta)P(t,T+\delta)}{P(t,T)} + \frac{1}{2}\zeta^2\right).$$

Thus given the prices of the caplets, Brace and Musiela are able to invert from the caplet pricing formula, for $t = 0.25i$, $i = 1,2, \ldots, 47$,

$$\zeta^2(t) = Var\left[\log P(t,t+\delta)\right] = \int_0^{t}\left|\int_s^{s+\delta}\underline{\tau}(t-s,u)du\right|^2 ds,$$

which are fitted using a cubic spline to give $\zeta^2(t)$ for all $t \geq 0$. The function $\zeta^2(t)$ is decomposed into the sum of two positive functions

$$\zeta^2(t) = \zeta_1^2(t) + \zeta_2^2(t)$$

where

$$\zeta_1^2(t) = \inf_{s\geq t}\zeta^2(s).$$

Brace and Musiela finally assume a two factor model with

$$\underline{\tau}(t,x) = \begin{bmatrix} \tau_1(x) \\ \tau_2(t+x)I_{[0,M]}(t) \end{bmatrix}$$

where M is the smallest real such that supp $\zeta_2^2 \subset [0,M]$ and $I_{[0,M]}(t) = 1$ for $0 \leq t \leq M$, and zero otherwise. The real functions τ_1 and τ_2 are extracted recursively using

$$\zeta_1^2(t) = \int_0^t \left(\int_s^{s+\delta} \tau_1(u)du \right)^2 ds$$

$$\zeta_2^2(t) = (t \wedge M) \left(\int_t^{t+\delta} \tau_2(u)du \right)^2$$

and the prices of the futures options.

The primary contribution made by Brace and Musiela (1994) on calibration is that they illustrate that it is not necessary to assume functional forms for the volatility factors $\underline{\tau}(t, x)$ and so they introduced non-parametric estimation of the volatility factors. This resolves the problem of determining an appropriate functional form highlighted in our review of Amin and Morton (1994) above. We feel however that their paper failed to address the following points:

1. The outlined method attempts to calibrate a two factor model using instruments that are insensitive to the correlation structure of zero coupon yield changes. To calibrate multifactor models, it is necessary to calibrate to options whose values are sensitive to the correlation structure of yield changes. Their calibration set does not included such instruments and they proceed to show how their calibrated model can be used to price and hedge swaptions. It is unlikely that the model swaption prices will be consistent with market prices.
2. A well specified and calibrated interest rate derivative pricing models would probably produce stable parameters between re-calibration through time. A model with wildly varying parameters between re-calibrations probably suffers from severe misspecification problems that suggest it is not capturing the interest rate dynamics well. If so, then the model is unlikely to price other types of derivatives accurately. Brace and Musiela (1994) do not show whether their fitted model has stable volatility functions between re-calibration through time.

We address these points in Chapter 7 where we calibrate a multifactor Gaussian HJM model to a wide range of caps and swaptions prices simultaneously.

3.3.3 Brace, Gatarek and Musiela (1995)

This paper introduces the market-Libor model and also shows how it can be calibrated to caps and swaptions[2] prices and a historically estimated correlation matrix of forward rate changes. We examine this model more carefully in Chapter 8 where we price resettable caps and floors, to be defined there, using the market-Libor model.

We avoid reproducing a summary of the model since it is presented in Chapter 8. We use a notation that is consistent with Chapter 8. Basically, the time t forward δ-Libor rate with maturity T for borrowing or lending over the period $[T, T+\delta]$, $L(t, T)$, can be modelled by the process

$$dL(t,T) = L(t,T)\underline{\lambda}(t,T)\cdot d\underline{W}_t^{T+\delta}$$

where $\underline{\lambda}(t, T)$ is a deterministic vector of the volatility factors and $d\underline{W}_t^{T+\delta}$ is a vector of Brownian increments with respect to the equivalent probability measure induced by taking the $T+\delta$ maturing PDB as the numéraire. This is consistent with the market quoting convention and this is why the model has become known as the market-Libor model. The lognormality of the forward δ-Libor rates allows some analytic tractability. Thus it is possible to price caps consistently with the market convention. Brace, Gatarek and Musiela (1995) also obtain an approximate

---

2 An interest rate swap is an agreement between two counterparties to exchange fixed interest payments for floating payments on an agreed principal for an agreed period. A payer swap allows the holder to pay fixed for floating and a receiver swap allows the holder to receive fixed for floating. The swap rates are arranged such that the counterparties can

swaption pricing formula. This allows them to calibrate their model to cap and swaptions prices and a historically estimated correlation matrix of forward rate changes much more quickly than otherwise possible. They assume that the deterministic Libor volatility structure is given by

$$\underline{\lambda}(t,T) = f(t)\begin{bmatrix} g_1(T-t) \\ g_2(T-t) \end{bmatrix}$$

where $g_1(\cdot)$ and $g_2(\cdot)$ are piecewise constant functions. If $f(\cdot) \equiv 1$, then the volatilities are time stationary. Brace, Gatarek and Musiela (1995) find that imposing the constraint $f(\cdot) \equiv 1$ does not allow their chosen volatility factors to fit options prices and correlations well but fits well otherwise.

Brace Gatarek and Musiela (1995) make a significant contribution by formalising the market-Libor model and showing how it can be calibrated. We feel their calibration is weak on the following areas:

1. $\underline{\lambda}(t, T)$ is not time-stationary, that is, it is not just a function of the maturity of the Libor rate. This means they have no control over the evolution of the volatility structure through time and cannot relate the implied volatility structure with historically measured values which are typically estimated assuming a time stationary volatility structure. Even if the correlations matrix is estimated allowing for time-dependence, the time dependence of $\underline{\lambda}(t, T)$ would still be unsatisfactory because it refers to a different time interval. Thus it is difficult to understand whether the implied volatility structure is realistic.
2. Their calibration only use swaptions with maturities up to two years which is short compared to the maturities available. Would adding longer maturity swaptions result in a much poorer fit?

---

enter the agreement without any costs. A swaption is an option that when exercised allows the holder to enter into a swap at a swap rate equal to strike of the option.

3. They do not test the stability of their calibrated volatility structure to provide some diagnostics on whether the model is likely to be poorly specified.

3.4 SUMMARY

We have noted that there are very few published papers on the calibration of interest rate term structure models to the interest rate term structure and options prices. We have reviewed a number of key papers. We highlighted their shortcomings and indicated how the remaining chapters in this thesis are related to them.

3.5 REFERENCES

1) Amin K. and Morton A.J., "*Implied Volatility Functions in Arbitrage-Free Term Structure Models*", Journal of Financial Economics, 1994, Vol. 35, pp 141-180.

2) Black F and Karasinski P, "*Bond and Option Pricing when Short Rates are Lognormal*", Financial Analysts Journal, July-Aug 1991, pp 52-59.

3) Black F, E Derman and W Toy, "*A One-Factor Model of Interest Rates and It's Application to Treasury Bond Options*", Financial Analysts Journal, Jan-Feb 1990, pp 33-39.

4) Brace A. and Musiela M., "*A Multifactor Gauss Markov Implementation of Heath, Jarrow and Morton*", Mathematical Finance, July 1994, Vol. 4(3), pp 259-283.

5) Brace A., Gatarek D. and Musiela M., "*The Market Model of Interest Rate Dynamics*", Working Paper, May 1995, University of New South Wales.

6) Chan K.C., Karolyi G.A., Longstaff F.A., and Sanders A.B., "*An Empirical Comparison of Alternative Models of the Short Term Interest Rate*", Journal of Finance, Vol. XLVII, No. 3, pp 1209-1227.

7) Cox J.C., Ross S.A. and Rubinstein M., "*Option Pricing: A Simplified Approach*", Journal of Financial Economics, 1979, pp 229-263.

8) Gibbons M.R. and Ramaswamy K, "*A Test of the Cox, Ingersoll and ross Model of the Term Structure*", Review of Financial Studies, Vol. 6(3), 1993, pp 619-58.

9) Heath D., Jarrow R., and Morton A.., "*Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation*", Econometrica, Vol. 60, No. 1, 1992.

10) Heath D., Jarrow R.A. and Morton A., "*Contingent Claim Valuation with a Random Evolution of Interest Rates*", Journal of Futures Markets, 1991, Vol. 10, pp 55-78.

11) Hull J. and White A., "*One Factor Interest Rate Models and the Valuation of Interest Rate Derivative Securities*", *Journal of Financial and Quantitative Analysis*, Vol. 28 (2), June 1993a, pp 235-254.

12) Hull J. and White A., "'*A Note on the Models of Hull and White for Pricing Options on the Term Structure': Response*", The Journal of Fixed Income, September 1995, pp 97-102.

13) Hull J. and White A., "*Efficient Procedures for Valuing European and American Path-Dependent Options*", The Journal of Derivatives, Fall 1993, pp 21-31.

14) Hull J. and White A., "*Using Hull-White Interest Rate Trees*", The Journal of Derivatives, Spring 1996, pp 26-36.

4. CALIBRATION OF SHORT RATE MODELS

4.1 INTRODUCTION

In this chapter we begin our examination of calibration by examining the calibration of short rate models. Short rate models were defined in Chapter 2. The short rate models use one time-dependent parameter to fit the initial interest rate term structure accurately. We argue a second time-dependent parameter has to be introduced to enable one factor short rate models to be calibrated accurately to options prices for those applications where good fits to options prices are important. We show how to fit the second time-dependent parameter to options prices. We provide simple examples and analyse the fitted models and their misspecification problems. We also discuss the calibration of multifactor short rate models.

There are various ways to calibrate short rate models to the term structures. In some cases we can express the short rate process parameters as functions of the observed initial term structures. Here we focus on short rate tree construction techniques that are useful for calibration and for pricing a wide range of interest rate derivatives. Short rate trees provide discrete time and space approximations for the distribution of the short rates which have processes that may feature time-dependent parameters.

We review the popular techniques pioneered in Jamshidian (1991) and Hull and White (1993a) for constructing short rate trees consistently with both the initial interest rate and volatility term structures in Section 4.2. In Section 4.3, we show how the short rate trees can be constructed consistently with options prices. We explain why for some applications, it is essential that the model reproduce market prices very accurately. We explain why the fitted models are generally badly misspecified and suggest that short rate models should be used to price

complex interest rate derivatives only when alternative methods are not preferred. We analyse the calibrated models and their misspecification. Section 4.4 discusses the calibration difficulties posed by multifactor short rate models. Section 4.5 summarises.

4.2 CALIBRATION OF SHORT RATE MODELS

As we described in the pricing review of Chapter 2, a short rate tree allows us to approximate the short rate distribution and value derivatives. In this section we provide a brief description of Black Derman and Toy (1990), (BDT), and Hull and White (1993a), (HW93), trees and their construction. BDT was originally presented in discrete form. Jamshidian (1991) shows that the continuous time limit of BDT is given by

$$d\ln r(t) = \left[\theta(t) - \frac{\sigma'(t)}{\sigma(t)}\ln r(t)\right]dt + \sigma(t)d\widetilde{W}(t)\,. \tag{4.1}$$

We examine HW93 models of the form

$$dr(t) = \left[\theta(t) - \phi(t)r(t)\right]dt + \sigma(t)r(t)^{\beta}d\widetilde{W}(t)\,. \tag{4.2}$$

These short rate models can be calibrated to fit an initial interest rate only or to fit an additional initial volatility term structure as well.

To fit the models to the interest rate term structure only, only one time-dependent parameter is required. Intuitively, making one time-dependent parameter provides the necessary degrees of freedom to fit market PDB prices. For BDT, $\sigma(t)$ is assumed to be constant with θ allowed to be time dependent. For HW93, both $\phi(t)$ and $\sigma(t)$ are assumed to be constant with θ allowed to be time dependent. In practice, β usually takes values 0 or 0.5 to give what are sometimes called Extended Vasicek and Extended CIR respectively. Chan, Karolyi, Longstaff and Sanders (1991), in a study using one-month Treasury-bill yields as a proxy for the short rate, provide an unconstrained estimate of 1.49 for β. It is now believed

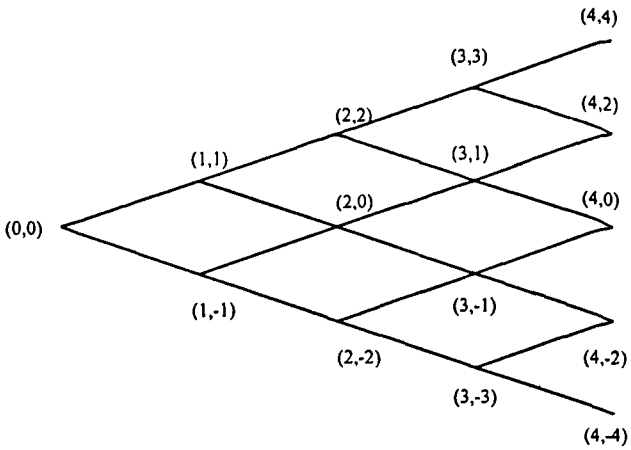
that their value of 1.49 is sensitive to outliers in their data. Furthermore, with a β value of 1.49, the short rates will explode with positive probability in finite time. To fit both term structures, additional parameters are provided by allowing an additional time-dependent parameter.

We shall see that a key feature in the construction of all short rate trees is the use of forward induction to grow the tree outward in time consistently with both the interest rate and volatility term structures. We will need Arrow-Debreu Pure Security Prices for the forward induction. We provide a short introduction to Arrow-Debreu Pure Security Prices before we describe the forward induction for the BDT and HW93 trees.

4.2.1 Arrow-Debreu Pure Security Prices, Kolmogorov Forward and Backward Equations

We first define some notation to identify the different tree nodes on a short rate tree. We assume that we will want a discrete approximation to the short rate distribution at time intervals of Δt. We show how this restriction can be relaxed in Section 4.3.1.1. The tree will be constructed with time steps of Δt and the short rate will be assumed to be the Δt zero-coupon yield.

The ith layer gives the attainable short rates at time $i\Delta t$. There are a finite number of attainable short rates at time $i\Delta t$ that are identified using an index j. We use the notation (i, j) to denote the state corresponding to $i\Delta t$ out in time with short rate $r(i\Delta t) = r(i, j)$. The current state is denoted by (0, 0). For example, in a recombining binomial tree, for the ith layer, the index j runs from $-i$ to i with increments of 2 so that there are $i + 1$ nodes in the ith layer:



Arrow-Debreu pure securities are contingent claims that have unit payoff when a particular state at a specified time is realised and have zero payoff otherwise. Clearly, the value of Arrow-Debreu securities depend on the current state and specified payoff state. We use $G_{n,i}[m, j]$ to denote the value of an Arrow-Debreu pure security that pays unit value at time $m\Delta t$ if $r(m\Delta t) = r(m, j)$ and nothing otherwise when the current state is (n,i).

We shall now proceed to derive a Kolmogorov Backward Equation for the Arrow-Debreu pure securities. We consider the case of a binomial recombining short rate tree. The derivation generalises to other cases. In a binomial tree, the node (n, i) moves to either $(n+1, i+1)$ or $(n+1, i-1)$. To remove unnecessary complicated notations for our discussion on the price relationship between Arrow-Debreu pure security prices, we also assume that node (n, i) moves to either $(n+1, i+1)$ or $(n+1, i-1)$ with equal risk neutral probability ½. Thus the value of the Arrow-Debreu security that pays unit value in the state (m, j), $G_{n,i}[m, j]$, moves to $G_{n+1,i+1}[m, j]$ and $G_{n+1,i-1}[m, j]$ with equal probability ½. Since Arrow-Debreu pure securities are contingent claims, then analogous to the pricing of European call options in the CRR binomial tree, their prices satisfy

$$G_{n,i}[m,j] = \frac{1}{2} dcf(n,i)\left\{G_{n+1,i+1}[m,j] + G_{n+1,i-1}[m,j]\right\} \qquad (4.3)$$

$$dcf(n,i) = \exp\{-r(n,i)\Delta t\}. \tag{4.4}$$

Equations 4.3 and 4.4 give the Kolmogorov Backward Equation for the prices of Arrow-Debreu pure securities in a binomial setting.

We will need the Kolmogorov Forward Equation to construct short rate trees. It will be clear from later discussions that the Kolmogorov Forward Equation is central to what is known as forward induction. The Kolmogorov Forward Equation and can be obtained as follows. Let G be the Arrow-Debreu pure security that pays unit value in state $(m+1, j)$ so that the price of G at state (m, k) is $G_{m,k}[m+1, j]$. Suppose we are in state (n, i) and consider a portfolio of consisting of $G_{m,k}[m+1, j]$ units of Arrow-Debreu pure security that pays unit value in state (m, k), for all k across the mth layer with cost

$$\sum_k G_{n,i}[m,k]G_{m,k}[m+1,j].$$

At time $n\Delta t$, the value of the portfolio will be just sufficient to buy G irrespective of the state we end up in. Therefore the portfolio allows us to duplicate the payoff to G and we must have

$$\sum_k G_{n,i}[m,k]G_{m,k}[m+1,j] = G_{n,i}[m+1,j]$$

which when combined with equation 4.3 gives

$$G_{n,i}[m+1,j] = \frac{1}{2}\left\{G_{n,i}[m,j-1]dcf(m,j-1) + G_{n,i}[m,j+1]dcf(m,j+1)\right\} \tag{4.5}$$

with $G_{n,i}[m, k] = 0$ for $k > m$ and $k < -m$. Equation 4.5 is the Kolmogorov Forward Equation for the prices of Arrow-Debreu pure securities in a binomial setting. Notice that the same arguments used for deriving equation 4.5 can be used to derive the Kolmogorov Forward Equation for the prices of Arrow-Debreu pure securities in more general settings. It is clear from the derivation that $G_{n,i}[m, j] = 0$ if it is impossible to move from state (n, i) to state (m, j).

We have introduced Arrow-Debreu pure securities and can now proceed to see how short rate trees can be constructed using forward induction.

4.2.2 BDT tree construction

We keep details to a minimum and refer the reader to BDT (1990) and Jamshidian (1991) for more details. We will first set up the tree.

The BDT tree is binomial. We use the same notation as in Section 4.2.1. That is, node (n, j) denote the state corresponding to time $n\Delta t$ out in time and when the short rate is at a level that is indexed by j. For the nth layer, j takes values from $\{-n, -n+2, \ldots, n-2, n\}$. The short rate applies to borrowing and lending over length Δt. Node (0, 0) corresponds to the initial state.

BDT assume that the short rates are lognormally distributed across each layer. The binomial tree allows the short rate at node (n, j) to move, with risk-neutral probability ½ and ½, to the short rates at nodes $(n+1, j+1)$ and $(n+1, j-1)$. We now show how the BDT tree can be fitted.

4.2.2.1 Fitting Black, Derman and Toy to the Interest Rate Term Structure Only

For the case where we are fitting to the interest rate term structure only, BDT assume that $r(n, j)$ is given by

$$r(n,j) = B(n)\exp[j\sigma\sqrt{\Delta t}]. \tag{4.6}$$

It is assumed that the initial short rate and its volatility is known. From equation 4.6 we can see that fitting in this case corresponds to solving for $B(i)$, for $i = 1, \ldots, N$, to make the constructed tree consistent with the initial interest rate term structure. N is the size of the binomial tree.

Rather than solving for $B(i)$, for $i = 1, \ldots, N$, simultaneously which would be extremely difficult, especially when the tree has to be large, Jamshidian (1991)

provides a technique that can be used to determine $B(\cdot)$ efficiently. Jamshidian (1991) introduces the forward induction technique that allows $B(\cdot)$ to be determined sequentially as the short rate tree is grown outwards step by step.

To see how the forward induction technique works. assume the short rate tree has been constructed consistently with the interest rate term structure up to the $(n\text{-}1)$th tree layer. To build another layer, we need to determine $B(n)$. $B(n)$ is related to the price of the $(n+1)\Delta t$ maturity PDB as follows.

Let $P(n+1)$ denote the time 0 value of a $(n+1)\Delta t$ maturity PDB that has to be fitted. $P(n+1)$ has value given by

$$P(n+1) = \sum_{j=-n}^{n} G_{0,0}[n,j]dcf(n,j) \tag{4.7}$$

$$dcf(n,j) = \exp\{-r(n,j)\Delta t\} \tag{4.8}$$

$$r(n,j) = B(n)\exp[j\sigma\sqrt{\Delta t}]. \tag{4.6}$$

Provided we have $G_{0,0}[n, j]$, for $j = -n, -n+2, \ldots, n\text{-}2, n$, equations 4.6 to 4.8 provide a non-linear equation for $B(n)$ that can be solved numerically. Here is where we use the Kolmogorov Forward Equation given by equation 4.5. We have assumed that the $(n\text{-}1)$st layer has already been constructed. The Arrow-Debreu pure security prices we need in equation 4.7 are given by the binomial forward equation

$$G_{0,0}[n,j] = \frac{1}{2}\left\{G_{0,0}[n-1,j-1]dcf(n-1,j-1) + G_{0,0}[n-1,j+1]dcf(n-1,j+1)\right\}. \tag{4.9}$$

Equation 4.9 follows easily from equation 4.5 with the appropriate substitutions. Thus equations 4.6 to 4.9 provide a single equation for $B(n)$ that can be solved numerically.

We have shown how the nth layer can be constructed given the $(n\text{-}1)$th layer. Thus if we can start the construction at the 0th layer, we can continue using

forward induction. The 0th layer is trivial with $G_{0,0}[0, 0] = 1$ and $r(0, 0) \equiv r_0$, the current value of the short rate. Thus the prescribed procedure allows the short rate tree to be grown layer by layer consistently with an initial interest rate term structure.

4.2.2.2 Fitting Black, Derman and Toy to both the Interest Rate and Volatility Term Structures

For the case where fitting to an initial volatility term structure and an initial interest rate term structure is desired, BDT assume the short rate at node (n, j) is given by

$$r(n,j) = B(n)\exp[j\sigma(n)\sqrt{\Delta t}]. \tag{4.10}$$

Comparing equation 4.10 with equation 4.6 shows that an additional time dependent parameter has been added to provide additional degrees of freedom to fit the short rate tree to a volatility term structure. In this case, we have to solve for $B(i)$ and $\sigma(i)$ for i = 1, ... , N, to make the constructed tree consistent with both the initial term structures.

To see how we can fit the tree, first observe that fitting the tree to an exogenously specified volatility term structure places constraints on how the zero coupon yields, or bond prices, move from node (0,0) to nodes (1,1) and (1,-1). Let $P_{1,1}(n+1)$ denote the value of the $(n+1)\Delta t$ maturity PDB at node (1,1) and let $P_{1,-1}(n+1)$ be defined similarly. Let $v(n)$ denote the return volatility on a $n\Delta t$ maturity PDB. The bond prices and the return volatility are related by

$$\log P_{1,1}(n+1) = \log P(n+1) + \alpha(n+1)\Delta t + v(n+1)\sqrt{\Delta t}$$

$$\log P_{1,-11}(n+1) = \log P(n+1) + \alpha(n+1)\Delta t - v(n+1)\sqrt{\Delta t}$$

where $\alpha(n+1)$ is the no-arbitrage drift rate. This pair of equations give, for each n, the following constraint on the PDB prices at (1, 1) and (-1, 1)

$$v(n+1) = -\sigma_Y(n+1)(n+1)\Delta t = \frac{1}{2\sqrt{\Delta t}} \log \frac{P_{1,1}(n+1)}{P_{1,-1}(n+1)} \tag{4.11}$$

where $\sigma_Y(n+1)$ is the absolute volatility of the $(n+1)\Delta t$ maturity zero coupon yield that we have to fit to. We also have to fit the initial bond prices. For each n, the initial bond prices provide another constraint

$$P(n+1) = \frac{1}{2} dcf(0,0)\left[P_{1,1}(n+1) + P_{1,-1}(n+1)\right]. \tag{4.12}$$

Equations 4.11 and 4.12 allow us to recover $P_{1,1}(n+1)$ and $P_{1,-1}(n+1)$ numerically. Thus fitting the BDT tree to an initial volatility term structure and an initial interest rate term structure involves finding $B(\cdot)$ and $\sigma(\cdot)$ such that the tree can reproduce the prices $P_{1,1}(\cdot)$ and $P_{1,-1}(\cdot)$.

We can now explain how to solve for $B(\cdot)$ and $\sigma(\cdot)$. Again, assume the $(n-1)$th layer has already been constructed. $B(n)$ and $\sigma(n)$ are related to $P_{1,1}(n+1)$ and $P_{1,-1}(n+1)$ as follows

$$P_{1,1}(n+1) = \sum_{j=-(n-1)}^{n} G_{1,1}[n,j] dcf(n,j) \tag{4.13}$$

$$P_{1,-1}(n+1) = \sum_{j=-n}^{n-1} G_{1,-1}[n,j] dcf(n,j). \tag{4.14}$$

$$dcf(n,j) = \exp\{-r(n,j)\Delta t\} \tag{4.8}$$

$$r(n,j) = B(n)\exp[j\sigma(n)\sqrt{\Delta t}]. \tag{4.10}$$

So given the Arrow-Debreu pure security prices $G_{1,1}[n, j]$ and $G_{1,-1}[n, j]$, equations 4.8, 4.10, 4.13 and 4.14 provide a pair of non-linear simultaneous equations for $B(n)$ and $\sigma(n)$ which can be solved numerically to complete the nth layer. Here is where we use the Kolmogorov Forward Equation given by equation 4.5. We have assumed that the $(n-1)$st layer has already been constructed. The Arrow-Debreu pure security prices we need in equations 4.13 and 4.14 are given by

$$G_{1,1}[n,j] = \frac{1}{2}\left\{G_{1,1}[n-1,j-1]dcf(n-1,j-1) + G_{1,1}[n-1,j+1]dcf(n-1,j+1)\right\} \quad \textbf{(4.15)}$$

$$G_{1,-1}[n,j] = \frac{1}{2}\left\{G_{1,-1}[n-1,j-1]dcf(n-1,j-1) + G_{1,-1}[n-1,j+1]dcf(n-1,j+1)\right\}. \quad \textbf{(4.16)}$$

Equations 4.15 and 4.16 follow easily from equation 4.5 with the appropriate substitutions.

We have shown how the nth layer can be constructed given the (n-1)th layer. We need to be able to start the tree. The first later is given by $dcf(1,1) = P_{1,1}(2)$, $dcf(1,-1) = P_{1,-1}(2)$, $G_{1,1}[1,1] = 1$ and $G_{1,-1}[1,-1] = 1$. Thus the tree can be grown layer by layer consistently with both term structures.

4.2.3 HW tree construction

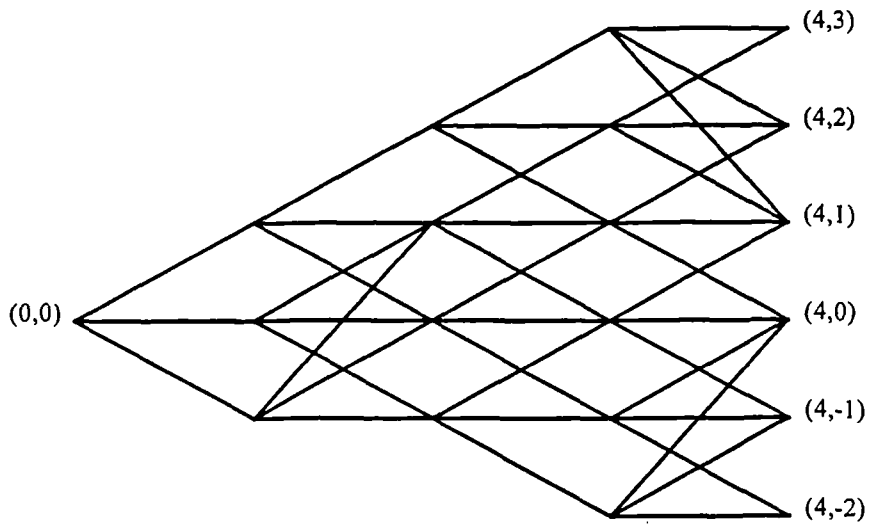
As for the BDT tree, we keep details to a minimum and refer the reader to HW(1993a) for full details. We illustrate the tree construction procedure for the case when $\beta = 0$. For the case when $\beta > 0$, HW93 shows a transformation of r can be used to generate a constant volatility process. The tree construction routine is then much the same. We can also use HW93 trees to implement Black and Karasinski (1991) by using the substitution $x = \ln r$ and constructing a tree for the process $dx = [\theta(t) - \varphi(t)x]dt + \sigma dz$.

HW93 trees are trinomial and the tree node (n, i) corresponds to the state $n\Delta t$ out in time with short rate $r_0 + i\Delta r$. That is

$$r(n,i) = r_0 + i\Delta r. \quad \textbf{(4.17)}$$

The short rate and Δt have the same interpretation as in the BDT tree and r_0 is the initial short rate. HW93 recommend $\Delta r = \sigma\sqrt{3\Delta t}$. This choice of Δr ensures that the branching probabilities remain positive and provides for good convergence as Δt decreases.

A key feature of the HW93 trinomial tree is that the branching structure changes to reflect the expected drift of the short rate so that the middle branch emanating from any node (n, i) reaches a node on the $(n+1)$st layer that is as close to the expected short rate $\tilde{E}\left[r(n+1)\middle|r(n)=r(n,i)\right]$ as possible given the short rate discretisation on the $(n+1)$st layer. For example, we may have a tree of the form



The construction techniques are similar to those of Sections 4.2.2.1 and 4.2.2.2. and we only sketch an outline.

4.2.3.1 Fitting Hull and White tree to Interest Rate Term Structure Only

To fit just an initial interest rate term structure HW93 assume $\phi(n) \equiv \phi$ so the short rate is assumed to follow the process $dr(t) = \left[\theta(t) - \phi r(t)\right]dt + \sigma d\tilde{W}(t)$. Fitting here corresponds to find $\theta(n)$ such that the tree agrees with an initial interest rate term structure.

To show how we can construct the tree, we first show how $\theta(n)$ is related to $P(n+2)$. We assume that the $(n-1)$st tree layer has already been constructed.

Let $P_{n,j}(n+2)$ denote the price of a $(n+2)\Delta t$ maturing PDB at the tree node (n, j). Then

$$P(n+2) = \sum_j G_{0,0}(n,j) P_{n,j}(n+2) \tag{4.18}$$

where $G_{0,0}[n,j]$ are Arrow-Debreu pure security prices as defined in Section 4.2.1. Note that $\theta(n\text{-}1)$ and ϕ are all that is needed to determine the branching structure from the $(n\text{-}1)$st layer to the nth later. Thus $\theta(n\text{-}1)$ will give the nth layer and the range of the summation in equation 4.18. The Arrow-Debreu pure securities, $G_{0,0}[n,j]$, are determined using a Kolmogorov forward equation similar to that of Section 4.2.1.

HW93 show we can relate $\theta(n)$ to $P(n+2)$ using equation 4.18 and an approximation for $P_{n,j}(n+2)$ given by

$$\begin{aligned} P_{n,j}(n+2) &= \exp\{-r(n,j)\Delta t\}\tilde{E}\left[\exp\{-r(n+1)\Delta t\}\middle| r(n) = r(n,j)\right] \\ &= \exp\{-2r(n,j)\Delta t\}\tilde{E}\left[\exp\{-\varepsilon(n,j)\Delta t\}\middle| r(n) = r(n,j)\right] \\ &\approx \exp\{-2r(n,j)\Delta t\}[1-(\theta(n)-\phi r(n,j))\Delta t^2] \end{aligned} \tag{4.19}$$

where $\varepsilon(n, j)$ is the random variable $\{r(n+1) - r(n) \mid r(n) = r(n, j)\}$. Equations 4.17, 4.18 and 4.19 give a closed form solution for $\theta(n)$. HW93 argue that the approximation provided by equation 4.19 introduces only very small errors and that the errors in the estimates for $\theta(\cdot)$ tend to be self correcting. HW93 argue that if the estimate for $\theta(n)$ is slightly too low then the estimate for $\theta(n+1)$ will be slightly too high and vice versa.

We now know how to extract $\theta(n)$ from $\theta(n\text{-}1)$ and $P(n+2)$. We can formalise the construction as follows.

Suppose the nth layer has already been constructed from $\theta(n\text{-}1)$. The $(n+1)$st tree layer can then be determined using the following three steps:

(1) Solve for $\theta(n)$ using equations 4.17, 4.18 and 4.19;

(2) Use $\theta(n)$ from step (1) to determine the branching structure from each node on the nth layer to the $(n+1)$st layer;

(3) Given the branching structure, determine the branching probabilities from each node in the nth layer to match the expected drift and variance rate. The probabilities are needed for determining $G_{0,0}[n+1,j]$ that are used in step (1) for the $(n+2)$nd layer.

These three steps allow us to extend the tree. We can construct the tree if we can start it. Node (0,0) is trivial and so the three steps allow the short rate tree to be grown outwards consistently with an initial interest rate term structure.

4.2.3.2 Fitting Hull and W tree to Interest Rate and Volatility Term Structures

To fit both the term structures, HW93 assume that the short rate process is $dr(t) = \left[\theta(t) - \phi(t)r(t)\right]dt + \sigma d\widetilde{W}(t)$.

To fit both the interest rate and volatility term structures, HW93 modify the branching structure so that node (0,0) has only two branches emanating from it: (0,0) moves to (1,1) and (1,-1) with equal risk neutral probabilities ½. All other nodes produce three branches. Similar to BDT above, fitting the short rate tree to both the term structures is the same as fitting to $P_{1,1}(\cdot)$ and $P_{1,-1}(\cdot)$ as defined in Section 4.2.2.2. Since $dcf(1,1) = P_{1,1}(2)$ and $dcf(1,-1) = P_{1,-1}(2)$, the short rates on the 1st layer do not have the form defined by equation 4.17. All other nodes do.

HW93 show how to choose $\theta(n)$ and $\phi(n)$ such that the tree is consistent with the term structures. To explain their technique, first observe that

$$P_{1,1}(n+2) = \sum_j G_{1,1}(n,j)P_{n,j}(n+2) \qquad (4.20)$$

$$P_{1,-1}(n+2) = \sum_j G_{1,-1}(n,j) P_{n,j}(n+2). \tag{4.21}$$

Similar to before, the range of the summations in equations 4.20 and 4.21 is determined by the branching structure from the (n-1)st layer to the nth layer that in turn depend on $\theta(n\text{-}1)$ and $\phi(n\text{-}1)$. Equations 4.20 and 4.21 need $P_{n,j}(n+2)$. HW93 show $P_{n,j}(n+2)$ can be approximated by

$$\begin{aligned} P_{n,j}(n+2) &= \exp\{-r(n,j)\Delta t\}\tilde{E}\left[\exp\{-r(n+1)\Delta t\}\middle| r(n) = r(n,j)\right] \\ &= \exp\{-2r(n,j)\Delta t\}\tilde{E}\left[\exp\{-\varepsilon(n,j)\Delta t\}\middle| r(n) = r(n,j)\right] \\ &\approx \exp\{-2r(n,j)\Delta t\}[1-(\theta(n)-\phi(n)r(n,j))\Delta t^2]. \end{aligned} \tag{4.22}$$

$\theta(n)$ and $\phi(n)$ can be solved analytically from equations 4.17, 4.20, 4.21 and 4.22 provided we have the Arrow-Debreu pure securities, $G_{1,1}[n+1, j]$ and $G_{1,-1}[n+1, j]$. The Arrow-Debreu pure securities can be determined using a Kolmogorov forward equation similar to that of Section 4.2.1.

We now know how to determine $\theta(n)$ and $\phi(n)$ from $\theta(n\text{-}1)$, $\phi(n\text{-}1)$, $P(n+2)$ and $\sigma_Y(n+2)$. We can now formalise the construction technique as follows.

Suppose the nth layer has already been constructed from $\theta(n\text{-}1)$ and $\phi(n\text{-}1)$. The (n+1)st tree layer is then determined using the following three steps:

(1) Solve for $\theta(n)$ and $\phi(n)$ using equations 4.17, 4.20, 4.21 and 4.22;

(2) Given $\theta(n)$ and $\phi(n)$ from step (1), determine the branching structure for the nth tree layer to the (n+1)st layer;

(3) Given the branching structure from step (2), determine the branching probabilities from each node in the nth layer to match the expected drift and variance rate. The probabilities are needed for determining $G_{1,1}[n+1, j]$ and $G_{1,-1}[n+1, j]$ needed in step 1 for the (n+2)nd layer.

These three steps allow us to extend the tree. We can construct the tree if we can start it. The first layer is also defined by $G_{1,1}[1,1] = 1$, $G_{1,-1}[1,-1] = 1$, $dcf(1,1) = P_{1,1}(2)$ and $dcf(1,-1) = P_{1,-1}(2)$. This also starts the forward induction so the three prescribed steps allow the short rate tree to be grown outwards consistently with both the term structures.

We have shown how BDT and HW93 short rate trees can be constructed to just an interest rate term structure or to both the term structures. Once the short rate tree has been constructed option prices follow easily. For some cases, such as Extended Vasicek, it is possible to price PDBs analytically. For others, the PDBs can be priced by taking expected discounted values back through the tree. In this way, the maturity values and early exercise values of many options can be found to allow the pricing of American / Bermudan / European options in a similar fashion to the Cox, Ross and Rubinstein (1979) binomial tree.

The calibrated trees here are, however, unlikely to be able to reproduce even the market prices of simple options. In all four cases, the estimation problems and misspecification problems discussed in Chapter 2 do not allow the models to price options consistently with the market. For consistency with market prices, rather than calibrating to historical volatilities, the models have to be calibrated to the market options prices.

4.3 CALIBRATION TO OPTIONS PRICES

We have argued that to be able to return consistent options prices with the market, the models have to be calibrated to options prices. Before we introduce our calibration, we will first describe a very simple calibration.

It is possible to calibrate short rate models to options prices by optimising the fit over the fixed parameters of the model. For example, in the HW93 model,

we may optimise the fit to market prices using $\phi(\cdot) \equiv \phi$ and $\sigma(\cdot) \equiv \sigma$; $\theta(\cdot)$, like before, is time dependent and fitted to the interest rate term structure. However, this calibration is unlikely to perform well because the volatility term structure is not flexible enough. The volatility term structure is just a function of two parameters. Furthermore, the fitted model may provide implausible values for the fixed parameters. For an accurate fit, the volatility term structure has to be more flexible. In this section, we show complete flexibility can be provided by allowing a second time dependent parameter.

4.3.1 Fitting to the Interest Rate Term Structure and Options Prices

In this section we show how we can calibrate a short rate tree to the zero coupon yield curve and options prices. As the short rate models considered here are one-factor models, we will only price simple options. We fit to caps and floor prices separately. The technique here can be extended in principle to multifactor models.

We will fit the trees to the prices of the caplets and floorlets[1]. So first consider how caplets are and floorlets are priced. Consider a caplet with maturity T, that pays $\delta[L(T) - k]^+$ at time $T+\delta$, where $L(T)$ is the spot δ-Libor rate at time T for borrowing or lending for a period δ and defined implicitly by

$$1 + \delta L(T) = 1 / P(T, T+\delta) .$$

Under the EMM induced by taking the $T+\delta$ maturity bond as the numéraire, the time 0 value of the caplet is given by

---

1 We assume we have extracted prices of individual caplets (floorlets) from quoted cap (floor) volatilities. Typically, this involves, as a first step, constructing a curve for the cap (floor) volatilities by interpolating between quoted cap (floor) volatilities. Then the price of the ith caplet (floorlet) is given by the difference between the prices of an $(i+1)/\delta$ maturity and an i/δ maturity cap (floor). In common with market practice, we assume that the first caplet (floorlet) starts in 3 months time.

$$cpt(0) = \delta P(0, T+\delta) E^{T+\delta}\left[L(T) - k\right]^+$$

where k is the cap rate. Changing to the risk-neutral measure and simplifying gives

$$cpt(0) = \tilde{E}\left[\exp\left\{-\int_0^T r(s)ds\right\}\left(1-(1+\delta k)P(T,T+\delta)\right)^+\right]. \tag{4.23}$$

Equation 4.23 shows that the value of the caplet is the same as a T maturity European put, strike 1, on a $T+\delta$ maturity PDB with face value $1+\delta k$. The price of the corresponding floorlet with payoff $\delta[\, k - L(T)]^+$ at time $T+\delta$, is the same as the pricing of the corresponding European call option on the PDB

$$flt(0) = \tilde{E}\left[\exp\left\{-\int_0^T r(s)ds\right\}\left((1+\delta k)P(T,T+\delta)-1\right)^+\right].$$

Thus we see that to price the caplet (floorlet) we need to determine the time T value of the $T+\delta$ maturity PDB from the short rate tree, (For the β_0 HW93 model, the PDB values are given analytically) and the short rate tree has to be constructed for a period δ beyond the caplet (floorlet) We will assume that the caplets (floorlets) have duration $\delta = 1/4$ year.

We need to set up the tree to price the caplets (floorlets). The step size has to be chosen to provide an integer number of steps within ¼ intervals. Let there be k tree steps per ¼ year. Then the ith caplet (floorlet) spans the tree layers ik to $(i+1)k$. To price the ith caplet (floorlet), we need to construct the short rate tree out to the $(i+1)k$th layer. To fit the n caplets (floorlets), we need $(n+1)k$ tree steps.

We describe how to fit HW93 trees to the caplet prices only. The procedure for the BDT tree is very similar. The procedure for floorlets is identical.

Except for the first three months, we assume we have end-to-end caplets (floorlet) that span the entire length of the short rate tree we need to construct.

Our calibration will solve for the implied volatility term structure directly rather than fitting the time dependent parameters directly. Given any volatility term structure we can fit the time dependent parameters are described in Section 4.2.

Let $\sigma_Y[j]$, the absolute volatility of the $j\Delta t$ zero coupon yield. We assume we have the short rate volatility and that zero coupon yield absolute volatilities are given by linear interpolations between the nodes $\sigma_Y[ik+1]$, $i = 2$ to $n+1$, and $\sigma_Y[0]$, which is the absolute volatility of the short rate:

$$\sigma_Y[j] = \sigma_Y[0] + \frac{j}{2k+1}\left\{\sigma_Y[2k+1] - \sigma_Y[0]\right\},\ 0 \le j \le 2k+1 \tag{4.24}$$

$$\sigma_Y[j] = \sigma_Y[ik+1] + \frac{j-(ik+1)}{k}\left\{\sigma_Y[(i+1)k+1] - \sigma_Y[ik+1]\right\}$$

$$\text{for } ik+1 \le j \le (i+1)k+1,\ 2 \le i \le n. \tag{4.25}$$

Our choice of volatility term structure follows from the following observations. There is no caplet covering the period $t \in [0, k\Delta t)$, so we cannot determine $\sigma_Y[j]$, $j =$ 1 to k, directly from the caplet prices. To fit the first caplet, the tree needs to be constructed up to the $2k$th layer and requires $\sigma_Y[2k+1]$. Similarly the 2nd caplet requires $\sigma_Y[3k+1]$ and so on. We have n parameters, $\sigma_Y[ik+1]$, $i = 2, \ldots ,.n+1$, to fit n caplet prices.

Our calibration fits the prices of the caplets sequentially. Therefore it is efficient. The first caplet is fitted by searching over $\sigma_Y[2k+1]$. For each trial value of $\sigma_Y[2k+1]$ we construct the tree out to the $2k$th tree layer and price the first caplet. We use the Newton-Raphson iterative scheme to determine the required volatility.

Now suppose the first j caplets have been priced consistently with the market so that the implied volatility term structure and the tree out to the $[(j+1)k]$th tree layer have been determined. We can extend the implied volatility

term structure by searching for $\sigma_Y[(j+2)k+1]$ until the $(j+1)$st caplet is priced consistently. For each trial value of $\sigma_Y[(j+2)k+1]$, the tree is constructed out to the $(j+2)k$th tree layer. The price of the first j caplets are unaffected by the volatility term structure beyond the $[(j+1)k+1]$th step and so the tree prices $j+1$ caplets consistently. In this way we can grow the grow the tree and propagate the implied volatility term structure outwards systematically to price all n caplets.

4.3.1.1 Example 1: Calibration to Caplets Prices

In this section, we fit the BDT tree and two HW93 trees, $\beta = 0$ and 0.5, to a zero coupon yield curve and caplet prices. Instead of fitting to market caps prices, we fit to theoretical prices to illustrate several important points that would not be so easy to present otherwise. We assume that the real world corresponds to an one-factor Gaussian HJM model and generate the required caplet prices. Chapter 7 shows how we can price caplets analytically in Gaussian models. We assume the zero coupon yield curve is 8.5% flat and the zero coupon yield volatility term structure is as plotted in Figure 1. The HJM volatility term structure resembles those produced by an examination of historical data. The volatilities fall off quickly as the maturities increase but flattens out to make it substantially different from an exponential decay. We fit BDT and HW93 models to the HJM caplet prices.

It is important to provide many tree steps up to the commencement of the first caplet to price it accurately. For this reason, the HW93 trees were constructed with $\Delta t = 1/40$ up to the end of the first caplet and with $\Delta t = 1/8$ thereafter. Note that the short rates at the 20th tree layer and beyond will be 1/8 year zero coupon yields whereas those before are 1/40 year zero coupon yields. To construct the tree beyond the 20th tree layer, it is necessary, for consistency, to determine the 1/8 year zero coupon yields for the 20th tree layer. We do this by extending the tree out by another 1/8 year, that is another 5 steps, to determine

the 1/8 year zero coupon yields. Having determined the 1/8 zero coupon yields, the tree construction method continues with $\Delta t = 1/8$. The BDT tree was constructed with 10 steps per ¼ year throughout the tree.

In practice, it is difficult to construct a short rate tree when the volatility term structure is not continuous. The branching structure can adopt unreasonable forms as a result. This is the reason we prefer a linear interpolation for the implied volatility term structure to a step function.

The implied volatilities are plotted in Figure 1 for comparison with the HJM input. We see from Figure 1 that the implied volatility term structures are substantially different from the assumed HJM input. Taking HJM and its term structure inputs as the real world, then BDT and HW93 models are misspecified and so cannot price the caplets consistently with HJM when using the HJM volatility input. Instead we can force the BDT and HW models to fit by bending the zero coupon yield volatilities. For caplets, calibration is achieved easily for all three models.

We examine the fitted model in more detail later. Here we note the example shows that even when interest rate models are misspecified, they may still allow calibration to some options prices. The value of the caplet considered in Section 4.3.1 is given by

$$cpt(0) = \delta P(0, T+\delta) E^{T+\delta}\left[L(T) - k\right]^{+}.$$

So intuitively, model calibration essential shifts the distribution of the Libor rates to allow the model to reproduce "market" prices.

Our next example will demonstrate that the ease with which a model can be calibrated to options depends on the model's ability to capture the key properties that give those options value.

4.3.1.2 Example 2: Calibration to Floorlet Prices

This example will illustrate the problems that can arise when calibrating a misspecified model. Consider a caplet with maturity T, that pays $\delta\,[k - L(T)]^+$ at time $T+\delta$, where k is the floor rate. The price of the floorlet can be obtained by evaluating

$$flt(0) = \delta P(0,T+\delta)E^{T+\delta}\left[k - L(T)\right]^+$$

directly or by using caplet-floorlet parity relationship

$$flt(0) = cpt(0) + (1+\tfrac{k}{4})\left(P(0,T+\tfrac{1}{4}) - \frac{P(0,T)}{1+\frac{k}{4}}\right)$$

where $cpt(0)$ is the price of the corresponding caplet.

When pricing caplets, the issue of whether to allow negative interest rates is unimportant. For floorlets, whether zero coupon yields can or cannot attain negative values can be extremely important. Suppose we again assume that the real world corresponds to a one-factor Gaussian HJM model where the zero coupon yield absolute volatility term structure is plotted in Figure 2. We assume the current zero coupon yield term structure yield curve is 2% flat and calibrate HW93 and BDT models to prices of ½% floorlets produced by the one-factor Gaussian HJM model.

What would the implied yield volatilities be for BDT and the HW models be? Whereas the Gaussian models allow negative rates and attribute value to the low struck floor, the non-negative models do not attribute much probability to low interest rates and would give little value to the floor. To reproduce the Gaussian HJM prices, the parameters for the non-negative short rate will have to be such the short rate mean-reverts very quickly to a very low, but positive, mean. In fact, our tree construction codes could not calibrate any of the non-negative short rate trees considered in this chapter to the low struck floorlet prices. If negative interest rates were possible in the real world, then non-negative models would be extremely

poor models for our example. Figure 2 shows the implied yield volatility for the HW93 model with $\beta = 0$.

Our two examples illustrate a fundamental modelling issue. Models should be tailored to an individual problem to ensure that the chosen model is capable of capturing the essential properties involved. For example, whether an interest rate model allows negative interest rates or not may or may not be important; it depends on the eventual use of the model.

We return to examine the fitted models. All the parameters are time dependent except for the short rate volatility parameter σ which is known since we assume we know the short rate volatility. The time-dependence compensates for model misspecification to allow model consistency with market prices. Figure 3 shows the fitted time dependent functions $\theta(n)$ and $\phi(n)$ for the HW93 model with $\beta = 0.5$ that was calibrated to the caplet prices. Figure 3 shows there are spikes in the fitted functions at the junctions between the linear segments of the implied volatility term structure. This suggests the parameters are sensitive to the second derivative of the volatility term structure. A deeper analysis is difficult because, in this case, it does not appear to be possible to express $\theta(t)$ and $\phi(t)$ as analytic functions of the current term structures. We have not plotted the fitted values for the other models since they behave in a similar fashion to those plotted in Figure 3.

We have shown how the short rate models can be calibrated to the term structure and cap or floor prices only. How well do these calibrated models price other contingent claims? We have assumed that our HJM model is the real world so that all the short rate models we have fitted are misspecified. They are different models and do not replicate the properties of the assumed HJM world. Our examples, provide a basis for examining real misspecification issues where no

models are capable of modelling the entire complexities of the real world accurately. In the next section we examine what calibration achieves and examine what problems arise from the differences between the short rate models and the assumed HJM world. We will see that this simple analysis will provide insights into real misspecification problems.

4.3.2 Misspecification Issues

To facilitate our analysis of misspecification problems, we will restrict our attention to the HW93 model with $\beta = 0$ because it provides some analytical tractability. We also suppose that the one factor Gaussian HJM model of sections 4.3.1.1 and 4.3.1.2 is the real world.

How is the HW93 model with $\beta = 0$ different from the one-factor Gaussian HJM model? Both are one factor models of the interest rate term structure with normally distributed zero coupon yields. The difference lies in the behaviour of the volatility term structure. If $v(t, s)$ denotes the volatility of the return on the PDB, $P(t, s)$, then it follows readily from Hull and White (1990) that the fitted HW93 model gives

$$v(t,s) = \frac{\sigma}{\partial v(0,t)/\partial t}[v(0,s) - v(0,t)], \; s > t, \tag{4.26}$$

where $v(0,\cdot)$ is the initial volatility term structure of PDB returns. Intuitively, as time progresses the volatility term structure will ride up the initial volatility curve with some scaling. The shape of the volatility term structure evolves through time in a deterministic manner. Carverhill (1995) shows the volatility structure will evolve through time in a deterministic manner except when the initial zero coupon yield volatilities are as in Vasicek (1977)

$$\sigma_Y(T) = \frac{\sigma}{\alpha T}(1 - e^{-\alpha T}). \tag{4.27}$$

However these negative exponential Vasicek volatility term structures do not fit empirical volatility term structures well. Furthermore, the HW93 model with $\beta = 0$ will only give the Vasicek volatility of equation 4.27 when $\phi(t) \equiv \phi$. The HW93 model with $\beta = 0$, when fitted to options prices allowing for both θ and ϕ to be time-dependent, is therefore very different from an one-factor Gaussian HJM model in general. The short rate processes implied by HJM models are in general non-Markovian. Carverhill (1994) shows that, with time-stationary volatility structures, the HJM implied short rate process will be Markovian only if the input volatility term structure is that of Vasicek (1977). The volatility term structure evolution of the HW93 model with $\beta = 0$ is thus a consequence of forcing the short rate process to be Markovian. This is a property of all Markovian short rate models when time-dependent parameters are introduced to fit options prices.

Now, historical data provide support for an evolving volatility term structure, but the calibrated trees may not capture the evolution in a suitable or accurate manner. This is a serious drawback of the BDT and HW93 models since many interest rate derivatives are sensitive to the future volatility term structure. Consider a callable bond that can be called at several different future dates. The volatility term structure at those different dates will affect whether the bond will be called. Having calibrated the HW93 model, there is no control on how the volatility term structure evolves through time. The HW93 model will therefore produce unreliable prices for the callable bond.

Because of these type of problems, Hull and White (1995, 1996) recommend that HW93 models should be provided with only one time-dependent parameter to allow consistency with an initial zero coupon yield term structure. However, without the extra time-dependent parameters, the models cannot be calibrated accurately to options prices. We have already argued that for some applications it is very important to reproduce each price in the calibration set accurately. For

example, consider an application where it is necessary to price and hedge a complex option with other simpler options. In this situation, the models will have to be calibrated accurately to the hedging options. This suggests that short rate models should be used only for short maturity options or for pricing instruments that would be very difficult to price with alternative approaches. For short maturity options, the volatility term structure will not have moved much and will therefore be similar to the initial implied volatility term structure and errors resulting from an unsuitably evolving volatility term structure will be limited. Short rate trees can be very useful for pricing very complex derivatives. For example, Hull and White (1993b) show how short rate trees can be extended to allow for the pricing of path-dependent interest rate derivatives.

It is very important that calibrated models can price and hedge options other than those used in the calibration. Practitioners typically re-calibrated their models through time to the then prevailing interest rate term structure and market option prices to update the hedge statistics. The updated parameters are generally inconsistent with previous estimates so that practitioners will in fact be using different models to hedge options through time but this problem would be unimportant if the re-calibration allows the model to hedge well. We want to assess how well misspecified models can hedge different options in practice. However this issue is difficult analyse directly and so we will investigate the hedging performance of a HW93 model with $\beta = 0$ calibrated to HJM caplet prices to provide insights.

Consider a caplet that covers the period $[k\delta, (k+1)\delta)$ with payoff $\delta[L(k\delta)-X]^+$ at time $(k+1)\delta$, where X is the cap rate and $L(k\delta)$ is the δ maturity Libor rate at time $k\delta$ given by the relationship $1+\delta L(k\delta) = 1/P(k\delta, (k+1)\delta)$. The value of this caplet is given by

$$cpt(0) = P(0,k\delta)N(-d_2) - (1+\delta X)P(0,(k+1)\delta)N(-d_1) \tag{4.28}$$

$$d_1 = \frac{1}{h}\log\frac{P(0,(k+1)\delta)}{(1+\delta X)P(0,k\delta)} + \frac{h}{2},\ d_2 = d_1 - h \tag{4.29}$$

$$h^2 = \int_0^{k\delta}\left\{v(u,(k+1)\delta) - v(u,k\delta)\right\}^2 du \tag{4.30}$$

where $v(u, w)$ is the return volatility on the PDB $P(u, w)$. Equations 4.28 to 4.30 are well known and we will provide their derivations in Chapter 7.

Now, in the case of HJM we have

$$\text{HJM: } v(u, w) \equiv v(0, w-u) \tag{4.31}$$

whereas for HW93 with $\beta = 0$ we have

$$\text{HW93 with } \beta = 0\text{: } v(u,w) = \frac{\sigma}{\partial\mu(0,u)/\partial u}[\mu(0,w) - \mu(0,u)] \tag{4.32}$$

where $\mu(0,\cdot)$ is the initial HW93 volatility term structure. Let $h_{HJM}(k\delta)$ denote equation 4.30 when equation 4.31 is used and $h_{HW93}(k\delta)$ to denote equation 4.30 when equation 4.32 is used. Then we can see that calibrating HW93 to the caplet prices is the same as finding the initial volatility $\mu(0,\cdot)$ that makes $h_{HW93}(k\delta) = h_{HJM}(k\delta)$ for all k. This observation provides important implications.

Firstly, even though the HW93 model is misspecified relative to the assumed HJM world, the HW93 model provides, instantaneously, the correct hedge ratios. To hedge a short position on the kth caplet over an instantaneous interval, long $N(-d_2)$ of the $k\delta$ maturing PDB and short $(1+\delta X)N(-d_1)$ of the $(k+1)\delta$ maturing PDB. Since both d_1 and d_2 for the calibrated HW93 model are the same as HJM, the hedge ratios are the same. This is, of course, only true for the instant when the HW93 model has been calibrated. If we allow continuous re-calibration through time, then the HW93 model will be able to hedge the caplets perfectly to maturity. Thus re-calibration can compensate for misspecification errors. Of course, in practice, it would not be feasible to re-calibrate and revise the hedge ratios

continuously but any hedging error resulting from misspecification would be minimal.

Secondly, note it can be readily shown that

$$h^2(k\delta) = k^2\delta^2 Var[Y(k\delta,(k+1)\delta)]$$

where $Y(k\delta, (k+1)\delta)$ is the δ maturity zero coupon yield at time $k\delta$, and is normally distributed by assumption. Therefore the calibration matches the distribution of the δ maturity zero coupon yields at times $k\delta$, $k = 1 .. n$, where n is the total number of caplets the HW93 model has been fitted to. The distribution of longer maturity zero coupon yields will not be matched and the calibrated HW93 model will even price European options on PDB differently from the assumed HJM world. To see this, consider a $k\delta$ maturity European call option on a $M\delta$ maturity PDB with strike X. The time 0 value of this European call option, $c(0)$, is well-known and is given by

$$c(0) = P(0,M\delta)N(d_1) - XP(0,k\delta)N(d_2) \tag{4.33}$$

$$d_1 = \frac{1}{h}\log\frac{P(0,M\delta)}{XP(0,k\delta)} + \frac{h}{2},\ d_2 = d_1 - h \tag{4.34}$$

$$h^2 = \int_0^{k\delta}\left\{v(u,M\delta) - v(u,k\delta)\right\}^2 du \tag{4.35}$$

where the PDB returns volatilities, $v(u,\cdot)$, are given by equation 4.31 for HJM and equation 4.32 for HW93. To see the difference, for the HJM and HW93 case respectively, equation 4.35 simplifies to

$$h^2_{HJM} = \int_0^{k\delta}\left\{v(0,(M-k)\delta + u) - v(0,u)\right\}^2 du \tag{4.36}$$

and

$$h^2_{HW93} = \left\{\sigma\left[\mu(0,M\delta) - \mu(0,k\delta)\right]\right\}^2 \int_0^{k\delta}\left\{\partial\mu(0,u)/\partial u\right\}^{-2} du \tag{4.37}$$

where $v(0,\cdot)$ is the time-stationary HJM PDB return volatility term structure and $\mu(0, \cdot)$ is the initial HW93 implied PDB return volatility term structure given by the calibration to the HJM caplet prices. For $M > k+1$, equations 4.36 and 4.37 will not agree except for perhaps coincidence. They have entirely different geometric interpretations. The calibrated HW93 model will therefore price the European call differently and give different hedge ratios from the assumed HJM world. This problem will increase with the maturity of the underlying zero coupon yield. We can conclude from this analysis that the calibrated HW93 model will reproduce what you calibrate to but give increasingly different results for non-calibrated derivatives as they diverge from the calibrated ones in what they depend on.

What conclusions can we draw from our analysis for the real world that no models are able to reproduce exactly? We can conclude that

1. The model must capture accurately the essential features of the real world that determine the value of the options that the model will be used to price;
2. The model must be re-calibrated through time to update the hedge ratios;
3. The calibration set must be similar to the eventual options that the model will be used to price.

4.4 MULTIFACTOR SHORT RATE MODELS

For the pricing and hedging of all but the simplest interest rate derivatives, it is essential that an interest rate derivatives pricing model captures the imperfectly correlation of zero coupon yield changes. However, single factor interest rate models imply all zero coupon yield changes are perfectly correlated. Thus single factor models may provide poor hedging performance and fail to price accurately. For example, single factor models would fail to price accurately spread options on the difference between two zero coupon yields.

There are various ways to produce multifactor models. One method is to generalise the above short rate models. For example, Langetieg (1980) and Longstaff and Schwartz (1992) produce multifactor short rate models by making the short rate an affine combination of stochastic variables. These models are however difficult to interpret and calibrate. For instance, in the two factor model $r = \alpha x + \beta y$, it is unclear what x and y represent and it is not transparent how α, β and other parameters in the stochastic processes for x and y affect zero coupon yield dynamics and options prices. Alternatively, we may follow Hull and White (1994) who produce a two-factor model that has the short rate reverting to a stochastic mean. Chen (1994) develops a three factor model with stochastic mean and stochastic volatility.

The multifactor short rate models can provide a wider range of shapes for the volatility term structure and a better fit to options prices than one factor models. However, the fitted parameters are sometimes implausible and when constrained to be within sensible ranges, the model may not fit so well. Therefore, it may still be necessary to provide an additional time-dependent parameter to ensure that the model prices a range of options consistently with the market. Thus the implied volatility term structure may again evolve unsuitably through time. Furthermore, it is far more difficult and time-consuming to construct multidimensional short rate trees. In view of these difficulties, multifactor versions of BDT, BK and HW93 short rate models may be unsuitable and impractical except for perhaps the pricing of some path-dependent options. We examine alternative models in the remaining chapters: Duffie and Kan (1995) provide a model where the state variables are zero coupon yields; we examine this model more closely in Chapter 5. HJM models allow for multiple factors readily; we examine HJM models and their calibration in Chapters 6, 7 and 8.

4.5 SUMMARY

We have reviewed how short rate models can be fitted to the interest rate term structure and volatility term structures simultaneously. We argued that the resulting models will not be able to price options consistently with the market and that the models should instead be calibrated to an interest rate term structure and options prices instead. Furthermore, for some applications, the models must price the calibration options accurately so that a second time-dependent parameter must be permitted. We provided simple examples to illustrate many modelling and calibration issues. We showed that

- whether an interest rate model allows negative rates or not is sometimes unimportant;
- whether an interest rate model is suitable or not depends on its intended use;
- interest rate models may exhibit unintended properties such as the volatility term structure evolution.

We pointed out, however, that introducing a second time-dependent parameter forces the volatility term structure to evolve deterministically. This is not a problem in itself, but the model user has no control on how the volatility term structure will evolve. In particular, the model user cannot prevent the volatility term structure evolving to undesirable and unrealistic shapes. This suggests that practitioners should only allow a second time-dependent parameter with care. However, without the second time-dependent parameter, the short rate models are unlikely to be able to fit options prices well. Without the second time-dependent parameter, the user can fit to options prices by varying the fixed parameters. However, this is unlikely to be satisfactory. For example, in the HW93 models there are only two other fixed parameters assuming that β has already been chosen. There are insufficient degrees of freedom to fit a large number of options prices well and the fit may imply implausible values for the

fixed parameters. If the user wanted to determine how to price complicated options and determine how to hedge them with simpler options or to assess their values relative to simpler options, then the model will have to be calibrated to the simpler options accurately. Our discussion points to a need for an alternative approach that allows flexibility for the specification of the volatility term structure. This is provided by the HJM approach which is will discuss in Chapters 6, 7, and 8.

We also examined misspecification problems. Our simple examples showed how re-calibration through time can compensate for model misspecification. It demonstrated the importance of choosing a calibration set of options that are similar to the options that the user wants to price. In particular, our analysis showed that calibrating to caps and floors would not provide good pricing to options whose payoffs depend on long maturity zero coupon yields. This provides support for the common practice of calibration to caps and swaptions prices.

4.6 REFERENCES

1) Black F and Karasinski P, "*Bond and Option Pricing when Short Rates are Lognormal*", Financial Analysts Journal, July-Aug 1991, pp 52-59.

2) Black F, E Derman and W Toy, "*A One-Factor Model of Interest Rates and It's Application to Treasury Bond Options*", Financial Analysts Journal, Jan-Feb 1990, pp 33-39

3) Carverhill A.P., "*Interest Rate Option Valuation Models: a Note about the Behaviour of the Volatility Structure*", Working Paper, 1992, Hong Kong University of Science and Technology.

4) Carverhill A.P., "*A Note on the Models of the Hull and White for Pricing Options on the Term Structure*", Journal of Fixed Income, September 1995.

5) Carverhill A.P., "*When is the Short Rate Markovian?*" Mathematical Finance, Vol. 4, 1994.

6) Chan K.C., Karolyi G.A., Longstaff F.A., and Sanders A.B., "*The Volatility of Short-Term Interest Rates: An Empirical Comparison of Alternative Models of the Term Structure of Interest Rates.*", Journal of Finance, Vol. 47, July 1992.

7) Chen L., "Stochastic Mean and Stochastic Volatility: A Three-Factor Model of the Term Structure and its Application in the Pricing of Interest Rate Derivatives", Working paper, Federal Reserve Board, October 1994.

8) Cox J.C., Ross S.A. and Rubinstein M., "*Option Pricing: A Simplified Approach*", Journal of Financial Economics, 1979, pp 229-263.

9) Duffie D. and Kan R., "*A Yield-Factor Model of Interest Rates*", Working Paper, Graduate School of Business, Stanford University, August 1995.

10) Hull J. and White A., "*One Factor Interest Rate Models and the Valuation of Interest Rate Derivative Securities*", *Journal of Financial and Quantitative Analysis*, Vol. 28 (2), June 1993a, pp 235-254.

11) Hull J. and White A., "*Pricing Interest Rate Derivative Securities*", The Review of Financial Studies, Vol. 3 (4), 1990.

12) Hull J. and White A., "'*A Note on the Models of Hull and White for Pricing Options on the Term Structure': Response*", The Journal of Fixed Income, September 1995, pp 97-102.

13) Hull J. and White A., "*Efficient Procedures for Valuing European and American Path-Dependent Options*", The Journal of Derivatives, Fall 1993b, pp 21-31.

Hull J. and White A., "'*A Note on the Models of Hull and White for Pricing Options on the Term Structure': Response*", The Journal of Fixed Income, September 1995, pp 97-102.

14) Hull J. and White A., "Numerical Procedures for Implementing term Structure Models II: Two-Factor Models", The Journal of Derivatives, Winter 1994, pp 37-48.

15) Hull J. and White A., "*Using Hull-White Interest Rate Trees*", The Journal of Derivatives, Spring 1996, pp 26-36.

16) Jamshidian F., "*Forward Induction and Construction of Yield Curve Diffusion Models*", Merrill Lynch working paper, 1991.

17) Langetieg T.C., "*A Multivariate Model of the Term Structure.*", Journal of Finance, March 1980, pp. 71-97.

18) Longstaff F.A., Schwartz E.D., "*Interest-Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model.*", Journal of Finance, September 1992, pp. 1259-1282.

Figure 1: Implied Absolute Zero Coupon Yield Volatility Term Structures (Caps)

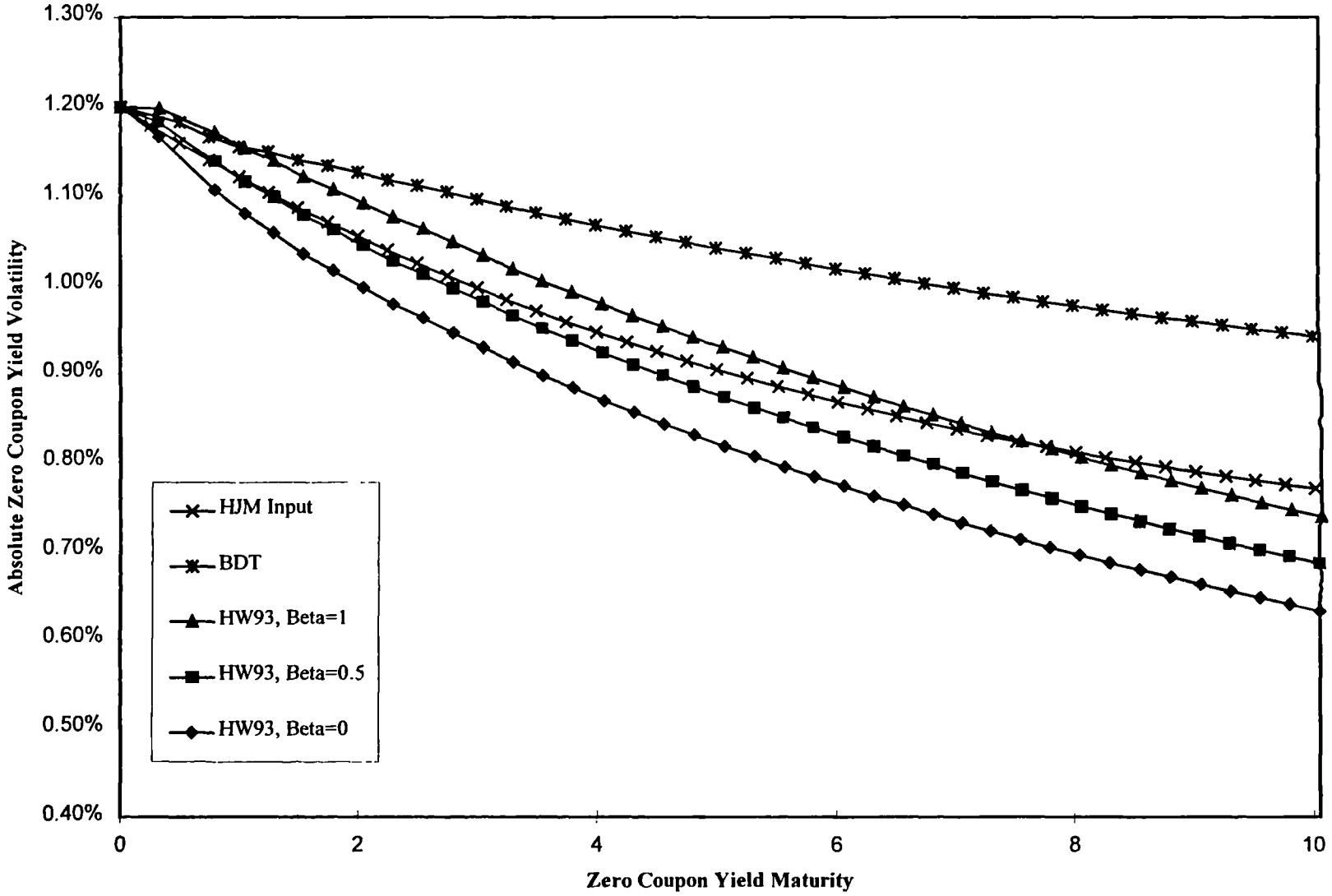


Figure 2: Implied Zero Coupon Yield Volatility Term Structure (Floors)

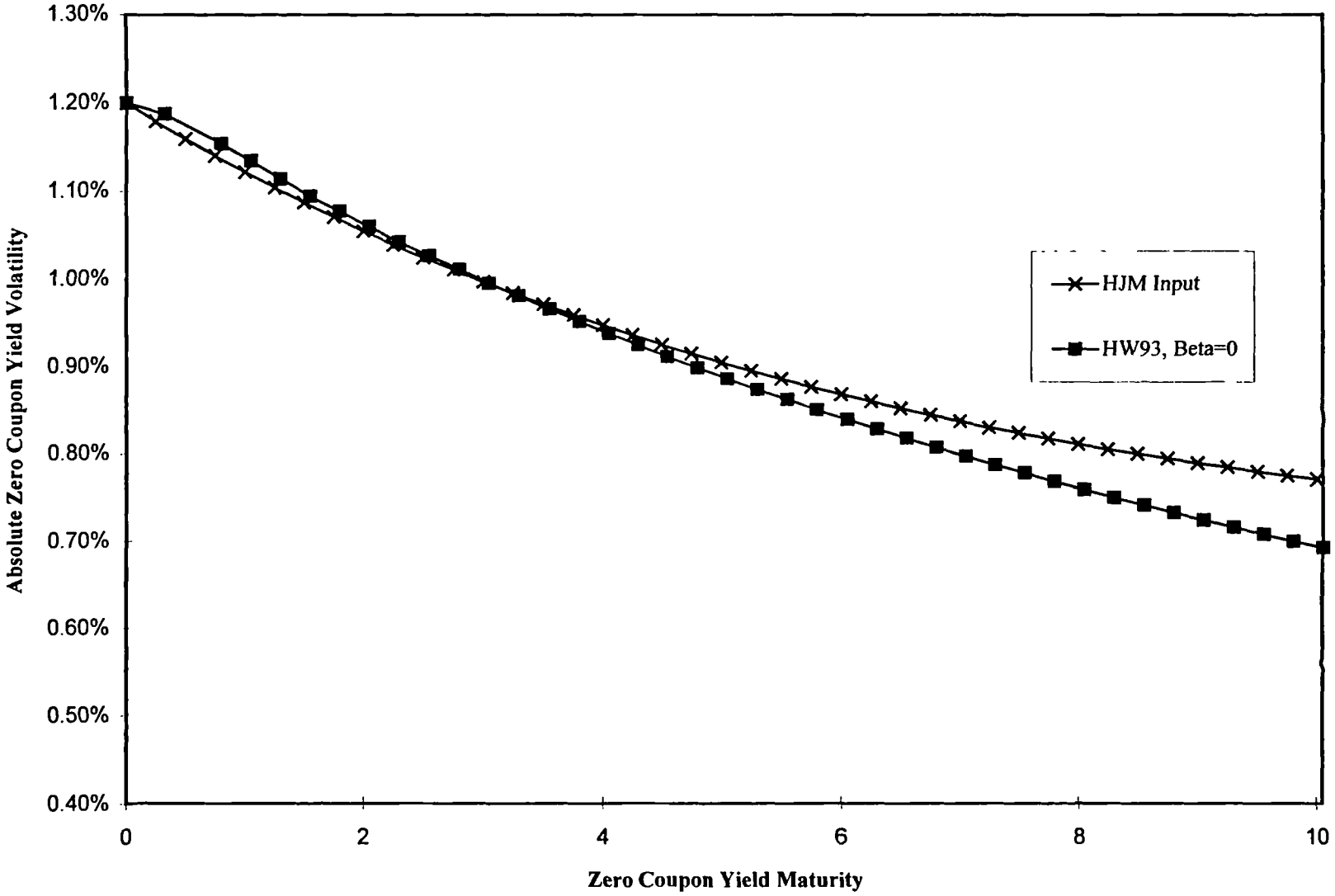
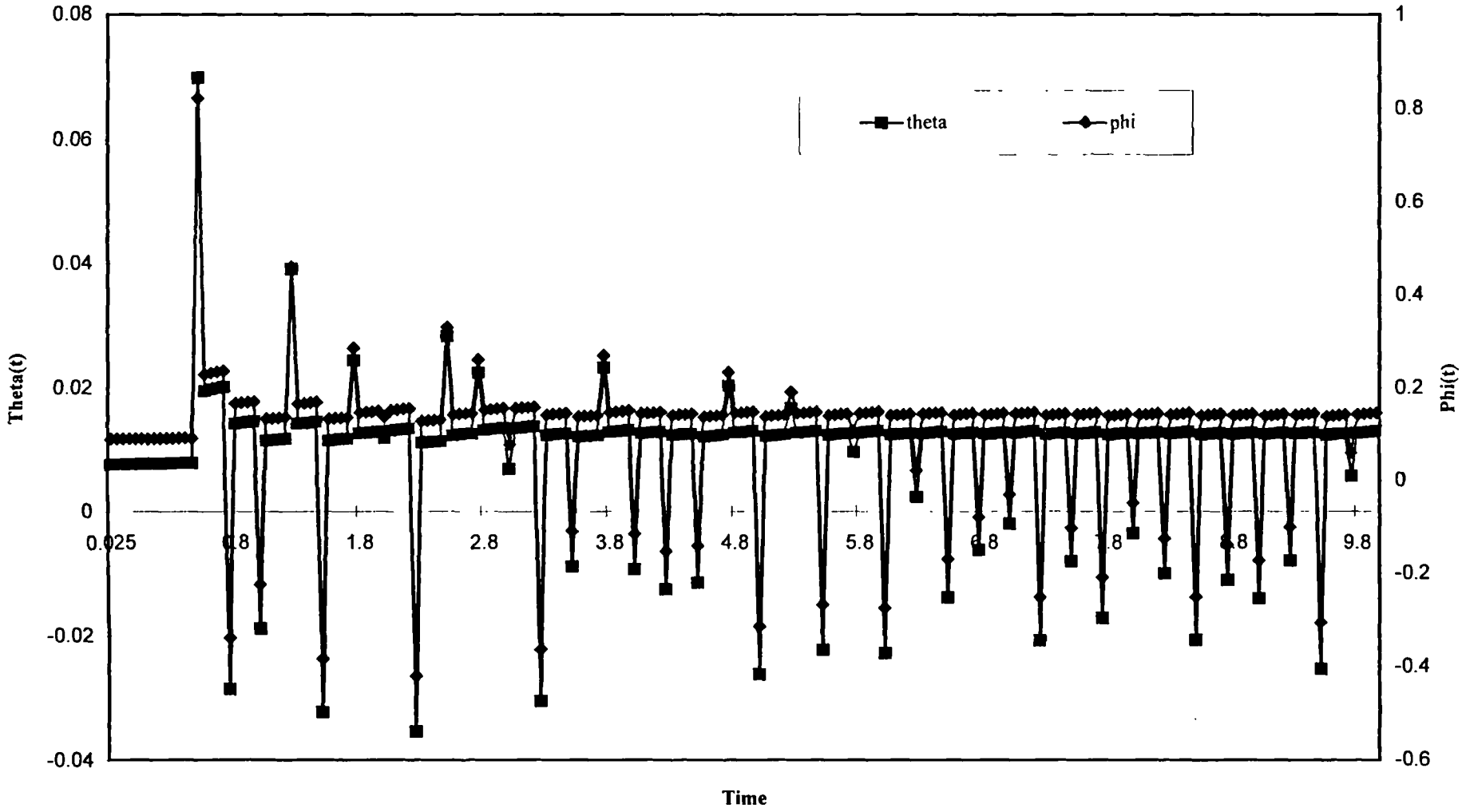


Figure 3: HW93, Beta=0.5, Fitted Theta and Phi



5. DUFFIE AND KAN AFFINE MODELS OF THE TERM STRUCTURE OF INTEREST RATES[1]

5.1 INTRODUCTION

We have seen in the previous chapter how single factor short rate models can be calibrated simultaneously to an interest rate term structure and options prices. Although calibration can be achieved with ease, the fitted single factor models are in many ways unsatisfactory. Perhaps the most serious problem is the non-stationarity of the implied volatility term structure when time-dependent parameters are introduced to ensure a good fit to options prices. This is particularly serious when the fitted tree is used to price more sophisticated options such as compound options where the volatility term structure affects the price of the underlying option greatly.

We also considered multifactor short rate models. The multifactor models can provide a wider range of range of time-stationary volatility term structures and can potentially fit options well without a second time-dependent parameter. However, it can be difficult to see how the different parameters affect interest rate dynamics and so difficult to develop intuition on how the parameters affect options prices. These problems make it difficult to assess the suitability of the model for pricing different options and difficult to use in options trading. Furthermore, multifactor short rate models also generally include difficult to measure quantities. Thus it is difficult to use historical data to provide a guide on the suitability of the fitted parameters.

---

[1] This chapter is based closely on Pang K. and Hodges S.D., "*Non-Negative Affine Yield Models*", FORC preprint 95/62, 1995, University of Warwick.

In this Chapter, we examine the affine yield factor models introduced by Duffie and Kan (1996)[2], (DK). The models are affine in the sense that bond yields[3] are affine functions of the state variables. DK present a consistent and arbitrage-free affine multifactor short rate model of the term structure of interest rates in which the state variables can be chosen to be yields of selected maturities that follow a parametric multivariate Markov diffusion process with stochastic volatility. Brown and Schaefer (1993) originated the idea of characterising the family of short rate processes consistent with an affine term structure model. DK generalised Brown and Schaefer (1993) to an arbitrary finite number of state variables.

The Duffie and Kan Affine Model generalises several other published affine models, which we call the simple affine models. Examples of simple affine models in the literature include Vasicek (1977), Langetieg (1980), Cox, Ingersoll and Ross (1985), Longstaff and Schwartz (1992), Fong and Vasicek (1991), Chen and Scott (1992) and Chen (1994).

We refer to the Duffie and Kan model with arbitrary state variables as the Duffie and Kan Affine Model and their model with the state variables specified as yields as the Duffie and Kan Affine Yield Model.

DK argue that their affine yield models are more attractive than simple affine models. There are a number of reasons why this may be the case.

Firstly, since the state variables are yields which are observable and whose covariance structure can easily be measured or implied from market data, the calibration of the Duffie and Kan Affine Yield Models should be more transparent and easier than of the simple affine models. The simple affine models of Longstaff

---

[2] Duffie and Kan (1996) show how their results can be extended to jump-diffusion processes. We restrict our analysis to the diffusion model.

[3] Yields refer to zero coupon yields unless otherwise stated.

and Schwartz (1992) and Fong and Vasicek (1991), for example, have yields and volatilities that are functions of the short rate and volatility of the short rate which cannot be observed.

Secondly, yields are direct inputs to the Duffie and Kan Affine Yield Model, so changing model parameters only changes the covariance structure. This property again aids calibration as changing some parameters of the simple affine models changes both the yields and their covariance structure making the numerical search for the optimal parameters more difficult.

Thirdly, yields and their covariance structure are the natural variables to consider when pricing many derivatives such as caps and swaptions. Having yields as the factors not only makes the calibration simpler but also provides more intuition and insight into the valuation and hedging of interest rate derivatives. The Duffie and Kan Affine Yield Models certainly seem attractive in view of the problems of the short rate models we have discussed.

The advantages of the Duffie and Kan Affine Yield Model are counterbalanced by the fact that the parameters of the Duffie and Kan yields process must satisfy more constraints than those of the simple affine models. Some of these constraints are interior and terminal conditions on a system of Ricatti equations. In general we have to employ numerical techniques to find suitable parameters that satisfy those conditions. This is a serious problem and finding appropriate parameter values may be even more demanding than the calibration issues of the simple affine factor models highlighted earlier. Finding parameters consistent with non-negative yields is even more difficult.

This chapter argues that the difficulties just described for obtaining a Duffie and Kan Affine Yield Model will force practitioners to work with simple affine models or within the HJM framework. Chapter 7 will show that we can readily calibrate Gaussian HJM interest rate models. So if practitioners were to work with

Gaussian interest rate models, then they will probably always prefer to work with the HJM approach rather than a Gaussian Duffie and Kan Affine Model. It is possible to generate more general Duffie and Kan Affine models in which the short rate is non-Gaussian but also permits a negative short rate. They are unlikely to be of much interest. Thus if practitioners were to work with Duffie and Kan models, then they would be interested in those models that do not permit negative interest rates. We therefore focus our attention specifically on the Duffie and Kan Non-Negative Affine Yield Model.

This chapter establishes that the class of Duffie and Kan Non-Negative Affine Yield Models is equivalent to a class of simple affine models we call the Generalised Cox, Ingersoll and Ross (CIR) Interest Rate Models, for which parameters can be easily chosen to ensure non-negative yields. The equivalence result provides an alternative method for calibrating Duffie and Kan Affine Yield Models consistent with non-negative yields. We argue the alternative is preferred and that in practice nobody would calibrate Duffie and Kan Non-Negative Affine Yield Models directly. Thus the Duffie and Kan Affine Yield Model offers no calibration advantages.

We begin in Section 5.2 by showing how the Duffie and Kan class of models is related to models in the current literature, providing an introduction to the Duffie and Kan Affine Models and defining the Generalised CIR Interest Rate Models. Section 5.3 compares two approaches to calibrating a Duffie and Kan Non-Negative Affine Yield Model: An indirect method that calibrates a Generalised CIR Interest Rate Model before converting it to its Duffie and Kan Non-Negative Affine Yield Model equivalent and a direct approach that works with the yields process and short rate specification. Section 5.4 derives a number of results finishing with the equivalence between Duffie and Kan Non-Negative Affine Yield Models and

Generalised CIR Interest Rate Models. Section 5.5 summarises. We adopt the notation of DK throughout.

5.2 THE DUFFIE AND KAN AFFINE MODELS AND THE GENERALISED COX INGERSOLL AND ROSS INTEREST RATE MODEL

In this section we first relate the class of Duffie and Kan Affine Models to the current literature before proceeding to describe the equations which characterise them. We also define a strict subset of the Duffie and Kan Affine Model we call the Generalised CIR Interest Rate Model. Readers are referred to DK for more details.

The Duffie and Kan Affine class of models offers some tractability and is a generalisation of many models in the literature. The Duffie and Kan Affine class includes as special cases the well known models of Vasicek (1977), Langetieg (1980), Cox, Ingersoll and Ross (1985), Fong and Vasicek (1991), Longstaff and Schwartz (1992), Chen and Scott (1992) and Chen (1994). In fact, we do not know of any interest rate models with analytic bond pricing formulae that does not belong to the Duffie and Kan Affine class. The Duffie and Kan parameters can also be allowed to be time dependent so that the Duffie and Kan Affine class also includes extended versions of the above models. Allowing the parameters to be time-dependent enables all of the above models to fit the entire term structure of interest rates. However, interest rate derivative payoffs rarely depend on the entire term structure. Typically, the payoff is completely determined by a few yields. Thus it would not be necessary to model the entire term structure. The Duffie and Kan model allows us to model the yields that matter. Other yields are given by an affine function of the modelled yields.

Duffie and Kan models are a special subset of HJM models where a fixed number of bond yields follow the joint process defined in DK. El Karoui, Geman and Lacoste (1995) derive restrictions on the HJM volatility functions to allow the term structure of interest rates to be characterised by a finite number of interest

rate state variables when interest rates are normally distributed. The popular lognormal interest rate models of Black, Derman and Toy (1990) and Black and Karasinski (1991) do not belong to the Duffie and Kan Affine class and do not have analytic bond pricing formulae.

5.2.1 Duffie and Kan Affine Model

For the n state variables $\underline{X}$, DK show that to achieve a bond pricing formula of the form

$$P(\underline{X},\tau) = \exp[A(\tau) + \underline{B}(\tau)^T \underline{X}] \tag{5.1}$$

where $A(\tau)$ is a deterministic function, $\underline{B}(\tau)$ is a column vector of n deterministic functions, τ is the bond maturity and P the bond price then the risk-neutral process for $\underline{X}$, subject to a non-singularity condition in DK[4], takes the form

$$d\underline{X} = (a\underline{X} + \underline{b})dt + \Sigma \begin{bmatrix} \sqrt{v_1(\underline{X})} & & 0 \\ & \ddots & \\ 0 & & \sqrt{v_n(\underline{X})} \end{bmatrix} d\underline{\tilde{W}} \tag{5.2}$$

with

$$v_i = \alpha_i + \underline{\beta}_i^T \underline{X}$$

where a and Σ are (n x n) matrices, $\underline{b}$ and $\underline{\beta}_i$ are (n x 1) column vectors, $d\underline{\tilde{W}}$ is a (n x 1) column vector of independent Brownian increments, where the parameters satisfy for all i

A1) For all $\underline{x}$ such that $v_i(\underline{x}) = 0$, $\underline{u}_i^T \beta^T (a\underline{x} + \underline{b}) > \dfrac{\underline{u}_i^T \beta^T \Sigma\Sigma^T \beta \underline{u}_i}{2}$ and

A2) For all j, if $(\beta^T\Sigma)_{ij} \neq 0$, then $v_i = v_j$

to give $\underline{X}$ a unique solution and the superscript T denotes matrix transpose. Conditions A1 and A2 are proofed in DK. Intuitively, condition A1 provides for a

---

[4] The matrix C, Duffie and Kan (1996), equation 3.8, page 385, has to be non-singular.

sufficiently large drift for $dv_i(\underline{x})$ when $v_i(\underline{x})$ approaches zero and condition A2 prevents the correlation between $v_i(\underline{x})$ and $v_j(\underline{x})$ from taking $v_i(\underline{x})$ across zero. DK show that Σ has to be non-singular when none of the factors are scalings of each other. We assume this throughout.

We define the matrix β as the matrix with vector $\underline{\beta}_i$ as its ith column so that we can write

$$\underline{v} = \underline{\alpha} + \beta^T \underline{X}. \tag{5.3}$$

Specifying the short rate, r, as an affine function of the factors $\underline{X}$,

$$r = f + \underline{G}^T \underline{X} \tag{5.4}$$

completes the specification of the Duffie and Kan Affine Model. To provide a more compact notation we use

$$\underline{v}^D \equiv \begin{bmatrix} v_1 & & 0 \\ & \ddots & \\ 0 & & v_n \end{bmatrix} \text{ and } \sqrt{\underline{v}^D} \equiv \begin{bmatrix} \sqrt{v_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{v_n} \end{bmatrix}$$

henceforth. We use $\underline{u}_i$ to denote an unit column vector with zeros for all its elements except for the ith element which has the entry 1.

Given the process for $\underline{X}$ and the specification of the short rate we can solve for the functions $A(\tau)$ and $B_i(\tau)$ for i = 1 to n, perhaps numerically, subject to the initial conditions

$$A(0) = 0 \text{ and } B_i(0) = 0,\ i = 1 \text{ to } n. \tag{5.5}$$

For example, in a two factor model with $r = f + G_1X_1 + G_2X_2$ the functions $A(\tau)$ and $\underline{B}(\tau)$ are the solutions to the Ricatti equations

$$\frac{dB_1(\tau)}{d\tau} = -G_1 + a_{11}B_1(\tau) + a_{21}B_2(\tau)$$

$$+\frac{\beta_{11}}{2}[\Sigma_{11}B_1(\tau)+\Sigma_{21}B_2(\tau)]^2+\frac{\beta_{12}}{2}[\Sigma_{12}B_1(\tau)+\Sigma_{22}B_2(\tau)]^2 \qquad (5.6)$$

$$\frac{dB_2(\tau)}{d\tau}=-G_2+a_{12}B_1(\tau)+a_{22}B_2(\tau)$$

$$+\frac{\beta_{21}}{2}[\Sigma_{11}B_1(\tau)+\Sigma_{21}B_2(\tau)]^2+\frac{\beta_{22}}{2}[\Sigma_{12}B_1(\tau)+\Sigma_{22}B_2(\tau)]^2 \qquad (5.7)$$

$$\frac{dA(\tau)}{d\tau}=-f+b_1B_1(\tau)+b_2B_2(\tau)$$

$$+\frac{\alpha_1}{2}[\Sigma_{11}B_1(\tau)+\Sigma_{21}B_2(\tau)]^2+\frac{\alpha_2}{2}[\Sigma_{12}B_1(\tau)+\Sigma_{22}B_2(\tau)]^2 \qquad (5.8)$$

subject to the initial conditions $A(0) = B_1(0) = B_2(0) = 0$. Note that the specification of the short rate directly affects the solution to the Ricatti equations. That is f and $\underline{G}$ are parameters of the model. The system of ordinary differential equations is a standard initial value problem which can be solved numerically using, for example, the fourth-order Runge-Kutta method as described in Press *et al* (1988).

5.2.2 The Duffie and Kan Affine Yield Model

When the underlying state variables $\underline{X}$ are yields, in addition to providing a unique solution for the stochastic differential equation for $\underline{X}$, we must also choose parameters such that when we price a bond with a maturity τ_i that is the same as for one of the chosen yields, the bond pricing formula is given by $P(\underline{X},\tau_i)=\exp(-\tau_i X_i)$. The parameters must therefore be chosen such that the bond pricing formula meets the following constraints:

$$A(\tau_j)=B_i(\tau_j)=0,\ j\neq i,\ B_i(\tau_i)=-\tau_i \quad for\ i = 1\ to\ n. \qquad (5.9)$$

The general bond pricing formula, equation 5.1, will then reduce to the required form when pricing a bond with a maturity that is the same as one of the chosen reference yields. As the functions $A(\tau)$ and $B_i(\tau)$ for $i = 1$ to n are solutions to Ricatti equations with interior and terminal conditions, finding suitable parameters to

satisfy these conditions may be difficult. We may attempt to solve the system of ordinary differential equations by the shooting method as described in Press *et al* (1988).

Note that for fixed parameters, the Duffie and Kan Affine Yield Models will not be able to reproduce an entire yield curve. The model is fitted to a finite number of yields with distinct maturities and the portions of the yield curve between chosen reference maturities are an affine function of the reference yields. It is possible to obtain an exact fit by allowing time dependent parameters analogous to the models presented in Chapter 4. However, that would make the calibration of Duffie and Kan Affine Yield Models even more complicated and would be an unnecessary complication for the pricing of interest rate derivatives that usually depend only on a few bond yields. Consider the time t value of a European call option, with maturity s and strike X, on a coupon bond with payoffs c_i at times q_i, $i = 1$ to n, which is given by

$$c(t) = P(t,s)E_t^s\left[\sum_{i=1}^{n} c_i P(s,q_i) - X\right]^+$$

$$= P(t,s)E_t^s\left[\sum_{i=1}^{n} c_i \exp[-Y(s,q_i)(q_i - s)] - X\right]^+$$

where $Y(s, q_i)$ is the q_i - s maturity yield at time s. We only need to model the evolution of n bond yields and it is not necessary to model the entire term structure.

5.2.3 Duffie and Kan Non-Negative Affine Yield Model

Having non-negative zero coupon yields is equivalent to having a non-negative short rate. Given that $A(\tau)$ and $\underline{B}(\tau)$ satisfy equations 5.5 and 5.9, Section 5.4.2 shows that the short rate is given by

$$r = \left[\left. \frac{\partial \underline{B}(\tau)^T}{\partial \tau} \right|_{\tau=0} (\beta^T)^{-1} \underline{\alpha} - \left. \frac{\partial A(\tau)}{\partial \tau} \right|_{\tau=0} \right] - \left. \frac{\partial \underline{B}(\tau)^T}{\partial \tau} \right|_{\tau=0} (\beta^T)^{-1} \underline{v} \tag{5.10}$$

so that sufficient conditions for non-negative yields are[5]

$$\left[\left. \frac{\partial \underline{B}(\tau)^T}{\partial \tau} \right|_{\tau=0} (\beta^T)^{-1} \underline{\alpha} - \left. \frac{\partial A(\tau)}{\partial \tau} \right|_{\tau=0} \right] \geq 0 \tag{5.11}$$

$$- \left. \frac{\partial \underline{B}(\tau)^T}{\partial \tau} \right|_{\tau=0} (\beta^T)^{-1} \geq \underline{0}. \tag{5.12}$$

For the example considered in Section, 5.2.1, substituting equations 5.6 to 5.8 into equations 5.11 and 5.12 give, after some simplification, respectively

$$f \geq (G_1 \quad G_2)(\beta^T)^{-1} \underline{\alpha} \tag{5.13}$$

$$(G_1 \quad G_2)(\beta^T)^{-1} \geq \underline{0}. \tag{5.14}$$

We can see that these sufficient conditions, in general, involve non-linear constraints on the parameters. Given that we have argued it is already difficult to find parameters to ensure that the state variables are yields, it is therefore clear that it will be even more difficult to implement Duffie and Kan Yield Models with non-negative yields.

5.2.4 Converting a Duffie and Kan Affine Model to a Yield Model

DK show that we can transform the Duffie and Kan Affine Model $P(\underline{X}, \tau) = \exp[A(\tau) + \underline{B}(\tau)^T \underline{X}]$ where $\underline{X}$ are general state variables (which may or may not be yields) to the Duffie and Kan Affine Yield Model with yields of

[5] Equations 5.11 and 5.12 are actually sufficient for strictly positive yields since conditions A1 and A2 provide strictly positive $\underline{v}$.

maturities τ_j, $j = 1$ to n by the substitution $\underline{Y} = \underline{g} + h^T\underline{X}$ whenever the matrix h $= -\left[\frac{B_i(\tau_j)}{\tau_j}\right]_{i,j}$ is non-singular. The vector $\underline{g}$ is given by $-\left[\frac{A(\tau_j)}{\tau_j}\right]_j$.

Proposition 1: *We can always find a set of distinct τ_j, $j = 1$ to n such that h is non-singular when the model is non-degenerate.*

By non-degenerate we mean that no variables are redundant in the sense that it is not possible to find a change of variables that allows the bond pricing formula to be specified with fewer factors. We show that if h is singular then the bond pricing formula has at least one redundancy.

Proof: Suppose h is singular for distinct non-zero choices of τ_j, $j = 1$ to n. Define

$$\mathrm{B} = \begin{bmatrix} B_1(\tau_1) & \cdots & B_1(\tau_n) \\ \vdots & & \vdots \\ B_n(\tau_1) & \cdots & B_n(\tau_n) \end{bmatrix}.$$

Then equation 5.1 implies

$$\begin{bmatrix} \frac{dP(\tau_1)}{P(\tau_1)} \\ \vdots \\ \frac{dP(\tau_n)}{P(\tau_n)} \end{bmatrix} = ()dt + \mathrm{B}^T d\underline{X} \tag{5.15}$$

where B is singular if and only if h is singular[6] when none of the chosen maturities is zero. As B is singular there exists a $\underline{\lambda}$ with at least two non-zero elements such that $\mathrm{B}\underline{\lambda} = \underline{0}$ since none of the columns of B is $\underline{0}$ when all the chosen reference yields are greater than zero. Define C to be identity matrix with a column replaced

---

[6] We have $h = -\begin{bmatrix} B_1(\tau_1) & \cdots & B_1(\tau_n) \\ \vdots & & \vdots \\ B_n(\tau_1) & \cdots & B_n(\tau_n) \end{bmatrix} \begin{bmatrix} 1/\tau_1 & & 0 \\ & \ddots & \\ 0 & & 1/\tau_n \end{bmatrix}$.

by $\underline{\lambda}$ such that C is non-singular. Then defining $E = B^T C$, $\underline{Y} = C^{-1}\underline{X} + \underline{D}$ and substituting in equation 5.15 gives

$$\begin{bmatrix} \dfrac{dP(\tau_1)}{P(\tau_1)} \\ \vdots \\ \dfrac{dP(\tau_n)}{P(\tau_n)} \end{bmatrix} = ()dt + Ed\underline{Y}. \tag{5.16}$$

The columns of the matrix E in equation 5.16 are the same as those of B except one which is $\underline{0}$ by construction. We have provided a change of variable such that the bond prices are now only driven by n-1 factors or less and therefore that the original formulation must have had at least one redundancy. **QED**.

We therefore assume there are no redundancies in the bond-pricing formula throughout this paper and that h therefore is always non-singular for distinct choices of maturities. We have shown that we can transform any Duffie and Kan Affine Model to a Duffie and Kan Affine Yield Model where the maturities are non-zero and distinct. We can also therefore convert to a set of distinct yields that includes the short rate as the short rate is an affine function of the other yields.

5.2.5 The Generalised Cox, Ingersoll and Ross Interest Rate Model

The Generalised CIR Interest Rate Model is the model that is obtained when we specify the short rate as an affine function of $\underline{Y}$ that follows the process of DK with the restriction that the matrix Σ be the identity matrix and $v_i(\underline{Y}) = Y_i$. That is

$$d\underline{Y} = (a\underline{Y} + \underline{b})dt + \sqrt{\underline{Y}^D}\, d\underline{\widetilde{W}} \text{ and } r = f + \underline{G}^T\underline{Y}. \tag{5.17}$$

We say the process $d\underline{Y}$ has the CIR property. This is a generalisation of the 2-factor CIR models of Longstaff and Schwartz (1992) and Chen and Scott (1992). It allows for an arbitrary n factors with interactions between their drifts. To obtain the bond pricing formula we solve for the functions $A(\tau)$ and $B_i(\tau)$ for i = 1 to n as before. It is simple to guarantee that the yields are non-negative and we consider

non-negative yields a defining property of the Generalised CIR Interest Rate Model. That is, all Generalised CIR Interest Rate Model have non-negative yields.

Proposition 2: *The short rate is non-negative if*

i) *for all* $\underline{Y}$ *with* $y_i = 0$, $(a\underline{Y}+\underline{b})_i > ½$ *for* $i = 1$ *to* n;

ii) $\underline{G} \geq \underline{0}$ *and* $f \geq 0$.

Proof: Condition (i) ensures that $\underline{Y}$ cannot attain zero and follows from condition A1. Condition (ii) ensures that the short rate is a non-negative affine function of $\underline{Y}$. **QED**.

These conditions are easy satisfied. $\underline{Y}$ mean reverts to $-a^{-1}\underline{b}$ when a is negative definite. The Generalised CIR Interest Rate Model is clearly a special case of the Duffie and Kan Affine Yield Model.

We will go on to show that Generalised CIR Interest Rate Models are equivalent to Duffie and Kan Non-Negative Affine Yield Models. Let us, for now, assume that the equivalence holds so that we can consider the calibration issues first and understand the importance of the equivalence result.

5.3 IMPLEMENTATION AND CALIBRATION ISSUES

We distinguish between implementation and calibration issues. Implementation problems are the problems of how to get a model to value instruments and calibration problems are those we face when we want to find parameters that allow the model to match market data.

Suppose we want to calibrate a Duffie and Kan Non-Negative Affine Yield Model. In this section we consider two approaches for obtaining suitable parameters: an indirect method that first estimates a Generalised CIR Interest Rate Model before transforming to its Duffie and Kan equivalent and a direct approach that estimates the parameters for the yields process. We first examine an indirect approach and find that the implementation is simple but the calibration is difficult

because it is necessary to search over a large number of variables. We then look at a direct approach and find considerable implementation problems. It is difficult to choose parameters that are consistent with the factors being yields. Calibration would be relatively simple if there were a quick way to choose suitable consistent parameters. In both cases we assume we only have accurate estimates of the covariance structure between yields. We are fitting the risk-neutral processes and cannot therefore expect to have good estimates of the drift parameters from examining the objective behaviour of the yields.

Note that calibrating to yield covariances is easier than calibrating to options prices. Our example will show that calibrating to yield covariances is already so difficult that to calibrate to options prices would be too difficult for practical purposes.

5.3.1 The Indirect Approach

We first illustrate the ideas with the Longstaff and Schwartz (1992) two factor affine model before explaining how we may proceed when we do not have an explicit bond pricing formula. Both cases can be easily implemented to be consistent with non-negative yields for the constraints guaranteeing non-negativity are simple non-negativity and linear constraints on model parameters made clear in Proposition 2 of Section 5.2.5. Calibrating the models in both cases is more difficult.

Starting essentially from a specification of 2 CIR variables in their general equilibrium model

$$dx = (\gamma - \delta x)dt + \sqrt{x}dW_1 \text{ and } dy = (\eta - \xi y)dt + \sqrt{y}dW_2$$

Longstaff and Schwartz show that r, the short rate, and V, the instantaneous variance of the short rate, are given by

$$r = \alpha x + \beta y \text{ and } V = \alpha^2 x + \beta^2 y$$

where α and β are positive parameters of their model. They solve initially for a bond pricing formula in terms of x and y which they then through a change of variables express in terms of r and V. Longstaff and Schwartz have performed a change of factors rather like DK. Expressing the formula as a function of two different yields instead would give a Duffie and Kan Non-Negative Affine Yield Model.

Suppose we wish to establish a model with the instantaneous rate and the one year yield as the state variables and we want to fit the model to the market yields and covariance structure between the two chosen yields. Transforming from the factors (r, V) to the short rate, $Y(0)$, and the 1 year rate, $Y(1)$, gives

$$P(\underline{Y},\tau) = \exp\left[E(\tau) + \underline{F}(\tau)^T \underline{Y}\right]$$

where

$$Y = \begin{bmatrix} Y(0) \\ Y(1) \end{bmatrix},\ Y(0) \equiv r,\ \underline{F}(\tau) = \begin{bmatrix} F_0(\tau) \\ F_1(\tau) \end{bmatrix}$$

$$E(\tau) = 2\gamma \ln A(\tau) + 2\eta \ln B(\tau) + \kappa\tau - \frac{D(\tau)}{D(1)}\left[\kappa + 2\gamma \ln A(1) + 2\eta \ln B(1)\right]$$

$$F_0(\tau) = C(\tau) - \frac{D(\tau)C(1)}{D(1)},\ F_1(\tau) = -\frac{D(\tau)}{D(1)}$$

where $A(\tau)$, $B(\tau)$, $C(\tau)$ and $D(\tau)$ are the functions of Longstaff and Schwartz (1992) reproduced in Appendix 1 and $A(1)$, $B(1)$, $C(1)$ and $D(1)$ are the respective functions evaluated at $\tau = 1$. The yields process can be shown to be

$$d\underline{Y} = \left(R\underline{Y} + \underline{S}\right)dt + T\begin{pmatrix} \sqrt{v_0} & 0 \\ 0 & \sqrt{v_1} \end{pmatrix} d\underline{\tilde{W}} \text{ with}$$

$$R = \begin{bmatrix} R_{00} & R_{01} \\ R_{10} & R_{11} \end{bmatrix} \text{ where } R_{00},\ R_{01},\ R_{10} \text{ and } R_{11} \text{ are defined in Appendix 2}$$

$$S = \begin{bmatrix} S_0 \\ S_1 \end{bmatrix} \text{ where } S_0 \text{ and } S_1 \text{ are defined in Appendix 2}$$

$$T = \begin{bmatrix} \alpha & \beta \\ -\alpha[C(1)+\alpha D(1)] & -\beta[C(1)+\beta D(1)] \end{bmatrix} \text{ and}$$

$$\underline{v} = \underline{g} + h^T \underline{Y} \text{ with}$$

$$\underline{g} = \begin{bmatrix} g_0 \\ g_1 \end{bmatrix} \text{ where } h^T, g_0 \text{ and } g_1 \text{ are defined in Appendix 2.}$$

In this case, all model properties are explicit functions of the model parameters. We calibrate the model to best fit five targets, the short rate, one year rate, and their covariance structure, using six variables α, β, γ, δ, η, and ν where $\nu = \xi + \lambda$ and λ is a risk parameter. We could minimise a weighted sum of squared differences between model and market values using a numerical search routine. Note that it may not be possible to obtain an exact fit with the five target values even though we are searching over six parameters. For instance, it is difficult to obtain a small correlation between yields in the Longstaff and Schwartz (1992) model. Note also that we do not need to be able to observe x and y, the two CIR variables. Their values are implicitly defined by the yields we are fitting the model to.

For Generalised CIR Interest Rate Models where we do not have analytic bond pricing formulae we can easily solve numerically equations 5.6 to 5.8 which in this case simplify to

$$\frac{dB_1(\tau)}{d\tau} = -G_1 + a_{11}B_1(\tau) + a_{21}B_2(\tau) + \frac{1}{2}B_1(\tau)^2 \qquad \textbf{(5.18)}$$

$$\frac{dB_2(\tau)}{d\tau} = -G_2 + a_{12}B_1(\tau) + a_{22}B_2(\tau) + \frac{1}{2}B_2(\tau)^2 \qquad \textbf{(5.19)}$$

$$\frac{dA(\tau)}{d\tau} = -f + b_1B_1(\tau) + b_2B_2(\tau) \qquad \textbf{(5.20)}$$

subject to $A(0) = 0$ and $\underline{B}(0) = \underline{0}$ and the conditions of Proposition 2. Solving the ordinary differential equations 5.18 to 5.20 gives equation 5.1, where $\underline{X}$ are the

unobserved CIR state variables, which can be transformed to a formula in terms of bond yields $\underline{Y}$ giving

$$P(\underline{Y},\tau) = \exp\left[C(\tau) + \underline{D}(\tau)^T \underline{Y}\right]$$

where

$$\underline{Y} = \underline{g} + h^T \underline{X},\ h = -\left[\frac{B_i(\tau_j)}{\tau_j}\right]_{i,j} \text{ and } \underline{g} = -\left[\frac{A(\tau_j)}{\tau_j}\right]_j .$$

The unobserved CIR state variables are given by $\underline{X} = (h^T)^{-1}(\underline{Y}-\underline{g})$. We iterate the search routine, with each iteration solving the ordinary differential equations once, until we have a good fit to observed yields and their covariance structure. Note that changing the model parameters changes only the covariance structure and not the reference yields. The reference yields are inputs and changing the model parameters changes the $A(\tau)$ and $\underline{B}(\tau)$ functions: The unobserved variables are constrained to take on different values that give the same observed reference yields which may include the short rate. This property can make it easier to search for the best parameters.

We have shown that it is in principle possible to calibrate whether we have or have not got closed form solutions to the Ricatti equations. There is, however, a practical problem in that it is necessary to search over a large number of parameters which increases rapidly with the number of factors. There are $(n + 1)^2$ parameters in a n-factor model.

We have explained how we can calibrate all Duffie and Kan Non-Negative Affine Yield Models indirectly by our equivalence result. The next section shows that it is much harder to implement Duffie and Kan Non-Negative Affine Yield Models directly.

5.3.2 The Direct Approach

In this section we explain how we might calibrate Duffie and Kan Non-Negative Affine Yield Models directly. It is far more difficult to implement Duffie

and Kan Affine Yield Models directly than by the indirect method just explained. We illustrate the difficulties with a two factor example. More factors will increase the difficulties still further.

Suppose we are given two yields of different maturities, their volatilities and correlation that we wish to calibrate the model to. We need to determine the parameters for equations 5.2 - 5.4 such that $\underline{\nu} \geq \underline{0}$ and that the functions $A(\tau)$, $B_1(\tau)$ and $B_2(\tau)$ of equation 5.1 satisfy equations 5.5 and 5.9. The difficulties arise almost entirely from having to solve equations 5.6 - 5.8 subject to equation 5.9, the interior and terminal conditions of equation 5.6 - 5.8. Unless there are analytic solutions to this system of Ricatti equations, we shoot the functions forward, that is, we integrate the equations from τ equal to zero to the longest chosen reference maturity. Having completed the numerical integration we measure the discrepancies between the boundary conditions and the solution values. The discrepancy measure can be a weighted sum of squared differences. We must iterate the shooting to reduce the discrepancy measure to zero and only after the measure has been reduced to an acceptably small value do we calculate the covariance structure. Otherwise the bond pricing formula would not be consistent with having yields as the factors. However, the covariance structure produced may not match the observed structure well enough and we must then solve the Ricatti equations again for different parameters until we produce a close match to the observed covariance structure. A simple calibration procedure may consists of two nested iterations. The inner loop ensures that the functions satisfy the boundary conditions. The outer loop minimises the discrepancy between the model and target covariance structures. The procedure is clearly difficult. Imposing the constraints given by equations 5.11 and 5.12 to preclude negative yields makes the numerical search for consistent parameters even more demanding. We cannot neglect the extra constraints as the following example

illustrates. Consider the following parameters for the joint process followed by the 1 year and 12 year yields

$$A = \begin{bmatrix} -1.31193 & 3.55331 \\ -0.088328 & -0.138073 \end{bmatrix}, \quad B = \begin{bmatrix} -0.20058 \\ 0.024806 \end{bmatrix},$$

$$\Sigma = \begin{bmatrix} -0.038705 & 0.372175 \\ -0.008299 & 0.043145 \end{bmatrix}, \quad \alpha = \begin{bmatrix} 23.0104 \\ 2.27766 \end{bmatrix},$$

$$\beta = \begin{bmatrix} 30.4077 & 5.84922 \\ -262.299 & -27.2783 \end{bmatrix}$$

with $$r = 0.21608 + 1.98915X_1 - 3.25205X_2.$$

They give absolute one year and twelve year yield volatilities of 5.35% and 1.43% respectively and a correlation of 0.975 when the one year yield and twelve year yields are 2.8% and 6.03% respectively. These parameters, however, do not satisfy the sufficient non-negative constraints given by equations 5.11 and 5.12. We can see that the parameters permit negative yields more clearly when we convert to the equivalent Generalised CIR model:

$r = 0.6y - 0.05x$ with

$$dx = (0.05 - 0.55x)dt + \sqrt{x}d\widetilde{W}_1,$$

$$dy = (0.2 - 0.9y)dt + \sqrt{y}d\widetilde{W}_2$$

$x(0) = 1.107692$ and $y(0) = 0.158974$.

Both x and y can attain zero and since y can attain zero independently of x, the formula for the short rate shows clearly that the short rate and hence all other yields can attain negative values. This example demonstrates a drawback of the Duffie and Kan Affine Yield Model; it is difficult to understand how the processes are behaving from a quick examination of the parameters.

Our example has shown that it is usually rather difficult to implement a Duffie and Kan Non-Negative Affine Yield Model directly. The previous section demonstrated that we can calibrate Generalised CIR Interest Rate Models with

non-negative interest rates readily but the difficulties with calibrating short rate models presented in Chapter 4 applies. The Generalised CIR Interest Rate Models are, however, far easier to work with than their equivalent Duffie and Kan Non-Negative Affine Yield Models.

5.4 EQUIVALENCE BETWEEN GENERALIZED CIR INTEREST RATE AND DUFFIE AND KAN NON-NEGATIVE AFFINE YIELD MODELS

As a preliminary step, we show that in the Duffie and Kan Non-Negative Affine Yield Model the matrix β must be non-singular. This allows us to count the degrees of freedom present in the Duffie and Kan Non-Negative Yield Model and observe that it is the same as that in the Generalised CIR Interest Rate Model suggesting the two may be equivalent. The non-singularity of β allows us to transform uniquely from the Duffie and Kan reference yields to variables with the CIR property to establish the equivalence.

Proposition 3: *In the Duffie and Kan Non-Negative Affine Yield Model, the matrix β is non-singular.*

The intuition of the proof is as follows: The state space for the stochastic differential equations 5.2 and 5.3 is contained by the intersection of half spaces defined by $v_i \geq 0$ for $i = 1$ to n. For the state space to contain only non-negative $\underline{X}$, it must be the case that the state space is bounded by a number of non-parallel boundaries equal to or greater than the dimension of the hyperspace and this implies β must be non-singular.

Proof: Consider the equation $\left[\underline{\beta}_1 \cdots \underline{\beta}_n\right]^T \Delta\underline{x} = \underline{0}$ which only has non-degenerate solutions $\Delta\underline{x} \neq \underline{0}$ when β is singular. $\Delta\underline{x}$ is a vector that is parallel to all the boundaries. Suppose β is singular and $\hat{\underline{X}}$ is a feasible point, $\underline{v}(\hat{\underline{X}}) > \underline{0}$. It

follows that $\hat{\underline{X}} + k\Delta\underline{x}$ is also a feasible point since $d\underline{v} = k\beta^T\Delta\underline{x} = \underline{0}$. However $\hat{\underline{X}} + k\Delta\underline{x} > \underline{0}$ cannot hold for all k, that is not all yields can be non-negative when β is singular. Hence β must be non-singular in the Duffie and Kan Non-Negative Affine Yield Model. **QED**.

5.4.1 Degrees of Freedom

We first introduce a restriction to Duffie and Kan Affine Yield Model specification to remove some redundant parameters and then show that a Duffie and Kan Affine Yield Model has the same number of degrees of freedom as a Generalised CIR Interest Rate Model.

Consider equations 5.2 and 5.3 where both β and Σ are non-singular. DK shows that $\beta^T\Sigma$ must be a diagonal matrix as $v_i \neq v_j$ for $i \neq j$ when β is non-singular. The restriction is necessary for non-negative $\underline{v}$ and follows immediately from an examination of the process for $\underline{v}$.

Proposition 4: *We can scale uniquely the matrices β and Σ such that $\beta^T\Sigma = I$ without making any difference to the stochastic process.*

Proof: Suppose $\beta^T\Sigma = D$. Define $\bar{\Sigma} = \Sigma D$, $\bar{\underline{v}} = D^{-2}\underline{\alpha} + D^{-2}\beta^T\underline{X} = \bar{\underline{\alpha}} + \bar{\beta}^T\underline{X}$.

Then $d\underline{X} = (a\underline{X} + \underline{b})dt + \bar{\Sigma}\sqrt{\bar{v}^D}\,d\tilde{\underline{W}}$ with $\bar{\underline{\beta}}^T\bar{\Sigma} = I$ has covariance structure is given by $\bar{\Sigma}\sqrt{\bar{\underline{v}}^D}\sqrt{\bar{\underline{v}}^D}\bar{\Sigma}^T = \bar{\Sigma}\bar{v}^D\bar{\Sigma}^T = \Sigma D\bar{\underline{v}}^D D^T\Sigma^T = \Sigma\underline{v}^D\Sigma^T$ which is the same as before the scaling. Thus the stochastic process for $d\underline{X}$ is the same as before the scaling and we can always choose β and Σ such that $\beta^T\Sigma = I$ in the Duffie and Kan Affine Non-Negative Yield Model. **QED**.

Proposition 5: *The Duffie and Kan Non-Negative Affine Yield Model has the same degrees of freedom as the Generalised Cox Ingersoll and Ross Interest Rate Model.*

Proof: Assume the chosen yields do not include the short rate. There are $2n^2 + 2n$ parameters in the yields process. There are $n+1$ parameters for the specification of the short rate. The bond pricing formula, equation 5.1, must satisfy the following restrictions:

	τ				
	τ_1	----	τ_i	----	τ_n
$A(\tau)$	$A(\tau_1) = 0$	----	$A(\tau_i) = 0$	----	$A(\tau_n) = 0$
$B_1(\tau)$	$B_1(\tau_1) = -\tau_1$	----	$B_1(\tau_i) = 0$	----	$B_n(\tau_1) = 0$
:	:		:		:
$B_j(\tau)$	:		$B_j(\tau_i) = 0$ if $j \neq i$		:
			$B_j(\tau_j) = -\tau_j$		
:			:		:
$B_n(\tau)$	$B_n(\tau_1) = 0$	----	$B_n(\tau_i) = 0$	----	$B_n(\tau_n) = -\tau_n$

These give a total of $n(n + 1)$ internal boundary conditions on the set of Ricatti equations for $A(\tau)$ and $B_j(\tau)$ so that the number of independent parameters in the Duffie and Kan Affine Yield Model is $(n+1)^2$.

The generalised CIR Interest Rate Model, equation 5.17 also has $(n+1)^2$ parameters, although we usually restrict f to be zero. Thus the Duffie and Kan Affine Yield Model has the same degrees of freedom as the Generalised CIR Interest Rate Model[7]. **QED**.

---

[7] For the case when the short rate is included as one of the reference yields, indicated by τ_1, say, the restrictions on $A(\tau_1)$ and $\underline{B}(\tau_1)$ are redundant because these restrictions are now the same as the initial conditions for maturity equal to zero and so there are $n + 1$ fewer restrictions. There are also $n + 1$ fewer degrees of freedom as $r = f + \underline{G}^T\underline{X}$ becomes $r = X_1$ so that the total degrees of freedom is unaltered.

To show that the Generalised CIR Interest Rate Model and the Duffie and Kan Non-Negative Affine Yield Model are equivalent we only need to show a Duffie and Kan Non-Negative Affine Yield Model can be transformed uniquely to a Generalised CIR Interest Rate Model since Section 5.2.4 has already shown that we can convert uniquely from an Duffie and Kan Affine Model to a Duffie and Kan Affine Yield Model.

5.4.2 Equivalence

Proposition 6: *The Duffie and Kan Non-Negative Affine Yield Model and the Generalised Cox Ingersoll and Ross Interest Rate Models are equivalent.*

Proof: Starting from a Duffie and Kan Non-Negative Affine Yield Model, we can express the short rate as a sum of CIR variables using the transformation defined by equation 5.3. Eliminating $\underline{X}$ from the bond pricing formula, equation 5.1, gives equation 5.10

$$r = \left[\left. \frac{\partial \underline{B}(\tau)^T}{\partial \tau} \right|_{\tau=0} (\beta^T)^{-1} \underline{\alpha} - \left. \frac{\partial A(\tau)}{\partial \tau} \right|_{\tau=0} \right] - \left. \frac{\partial \underline{B}(\tau)^T}{\partial \tau} \right|_{\tau=0} (\beta^T)^{-1} \underline{v}$$

with

$$\begin{aligned} d\underline{v} &= (\beta^T a \underline{X} + \beta^T \underline{b})dt + \beta^T \Sigma \sqrt{\underline{v}^D} d\underline{\widetilde{W}} \\ &= \left[\beta^T a (\beta^T)^{-1} \underline{v} + (\beta^T \underline{b} - \beta^T a (\beta^T)^{-1} \underline{\alpha})\right] dt + \beta^T \Sigma \sqrt{\underline{v}^D} d\underline{\widetilde{W}} \\ &= \left(E\underline{v} + \underline{F}\right)dt + \sqrt{\underline{v}^D} d\underline{W} = \left(E\underline{v} + \underline{F}\right)dt + \sqrt{\underline{v}^D} d\underline{\widetilde{W}} \end{aligned}$$

and obvious substitutions. Section 5.4.1 has already shown that β and Σ can be scaled to make $\beta^T\Sigma$ equal to the identity matrix. The two equations define a Generalised CIR Interest Rate Model. Furthermore, this transformation must be unique up to a permutation of the indices. Any other affine transformation would not maintain the CIR property. We can therefore conclude that the two models are equivalent as Section 5.2.4 has already shown that we can transform uniquely in the opposite direction. **QED**.

5.5 SUMMARY

In this chapter we have shown that Duffie and Kan Affine Yield Models cannot be calibrated easily. We have shown Duffie and Kan Non-Negative Affine Yield Models are equivalent to Generalised Cox-Ingersoll-Ross Interest Rate Models. We have examined the implementation and calibration issues and argued that, when we want non-negative yield models, Generalised Cox Ingersoll and Ross Interest Rate Models would be far easier to work with empirically so that, in practice, Duffie and Kan Non-Negative Affine Yield Models would have to be obtained by converting non-negative short rate models. Thus the calibration of Duffie and Kan Non-Negative Affine Yield Models shares all the problems with the calibration of short rate models. Furthermore, if practitioners were prepared to use Gaussian models, then they would probably use Gaussian HJM models instead since Gaussian HJM models, as we show in Chapter 7, are simple to calibrate. Thus practitioners are unlikely to find Duffie and Kan Non-Negative Yield Models useful except for those would prefer to work with the yields formulation after they calibrated the equivalent Generalised Cox, Ingersoll and Ross models.

5.6 APPENDIX

Appendix 1: The Longstaff and Schwartz (1992) Bond Pricing Formula

Longstaff and Schwartz (1992) show that bond prices within their model economy, where r is the short rate and V is the instantaneous variance of the short rate, are given by

$$P(r,V,\tau) = A^{2\gamma}(\tau)B^{2\gamma}(\tau)\exp[\kappa\tau + C(\tau)r + D(\tau)V]$$

where

$$A(\tau) = \frac{2\Phi}{(\delta+\Phi)(\exp(\Phi\tau)-1)+2\Phi},$$

$$B(\tau) = \frac{2\Psi}{(\upsilon + \Psi)(\exp(\Psi\tau) - 1) + 2\Psi},$$

$$C(\tau) = \frac{\alpha\Phi(\exp(\Psi\tau) - 1)B(\tau) - \beta\Psi(\exp(\Phi\tau) - 1)A(\tau)}{\Phi\Psi(\beta - \alpha)},$$

$$D(\tau) = \frac{\Psi(\exp(\Phi\tau) - 1)A(\tau) - \Phi(\exp(\Psi\tau) - 1)B(\tau)}{\Phi\Psi(\beta - \alpha)},$$

and $\upsilon = \xi + \lambda$, $\Phi = \sqrt{2\alpha + \delta^2}$, $\Psi = \sqrt{2\beta + \upsilon^2}$, $\kappa = \gamma(\delta + \Phi) + \eta(\upsilon + \Psi)$.

Appendix 2: Duffie and Kan Formulation Of the Longstaff and Schwartz (1992) Model Using The Short Rate And The One Year Yield As State Variables

We show here the joint process for short rate, $Y(0)$, and one year yield, $Y(1)$, within the Longstaff and Schwartz (1992) model. The process is obtained by expressing $Y(1)$ as a function of V followed by substitution. We obtain

$$d\underline{Y} = (R\underline{Y} + \underline{S})dt + T\begin{pmatrix} \sqrt{v_0} & 0 \\ 0 & \sqrt{v_1} \end{pmatrix} d\underline{\widetilde{W}} \text{ with } \underline{v} = \underline{g} + h^T \underline{Y}$$

where

$$R_{00} = -\frac{D(1)(\beta\delta - \alpha\upsilon) - (\upsilon - \delta)C(1)}{(\beta - \alpha)D(1)}, \quad R_{01} = \frac{(\upsilon - \delta)}{(\beta - \alpha)D(1)},$$

$$R_{10} = -C(1)\left[-\frac{\beta\delta - \alpha\upsilon}{\beta - \alpha} + \frac{(\upsilon - \delta)C(1)}{(\beta - \alpha)D(1)}\right] - D(1)\left[-\frac{\alpha\beta(\delta - \upsilon)}{\beta - \alpha} + \frac{(\beta\upsilon - \alpha\delta)C(1)}{(\beta - \alpha)D(1)}\right],$$

$$R_{11} = -\frac{(\upsilon - \delta)C(1)}{(\beta - \alpha)D(1)} - \frac{\beta\upsilon - \alpha\delta}{\beta - \alpha},$$

$$S_0 = \alpha\gamma + \beta\eta - \frac{(\upsilon - \delta)(-\kappa - 2\gamma \ln A(1) - 2\eta \ln B(1)}{(\beta - \alpha)D(1)},$$

$$S_1 = -C(1)\left[\alpha\gamma + \beta\eta - \frac{(\upsilon - \delta)(-\kappa - 2\gamma \ln A(1) - 2\eta \ln B(1)}{(\beta - \alpha)D(1)}\right]$$

$$- D(1)\left[\alpha^2\gamma + \beta^2\eta - \frac{(\beta\upsilon - \alpha\delta)(-\kappa - 2\gamma \ln A(1) - 2\eta \ln B(1)}{(\beta - \alpha)D(1)}\right],$$

$$g_0 = \frac{\kappa + 2\gamma \ln A(1) + 2\eta \ln B(1))}{\alpha(\beta - \alpha)D(1)}, \quad g_1 = -\frac{\kappa + 2\gamma \ln A(1) + 2\eta \ln B(1))}{\beta(\beta - \alpha)D(1)},$$

$$h = \begin{bmatrix} \underline{h}_0 & \underline{h}_1 \end{bmatrix}, \quad \underline{h}_0 = \begin{bmatrix} \dfrac{C(1) + \beta D(1)}{\alpha(\beta - \alpha)D(1)} \\ \dfrac{1}{\alpha(\beta - \alpha)D(1)} \end{bmatrix}, \quad \underline{h}_1 = \begin{bmatrix} -\dfrac{C(1) + \alpha D(1)}{\beta(\beta - \alpha)D(1)} \\ -\dfrac{1}{\beta(\beta - \alpha)D(1)} \end{bmatrix}.$$

5.7 REFERENCES

1. Black F., Derman E., Toy W., "*A One-Factor Model of Interest rates and Its Application to Treasury Bond Options*", Financial Analysts Journal, Jan-Feb 1990, pp. 33-39.
2. Black F., Karasinski P., "*Bond and Option Pricing When Short Rates are Lognormal*", Financial Analysts Journal, July-Aug 1991, pp. 52-59.
3. Brown R. and Schaefer S., "*Interest Rate Volatility and the Term Structure of Interest Rates*", Philosophical Transactions of the Royal Society: Physical Sciences and Engineering, 1993, Vol. 347: pp563-576.
4. Chen L., "*Stochastic Mean and Stochastic Volatility: A Three-Factor Model of the Term Structure and its Application in Pricing in Interest Rate Derivatives*", Working Paper, Federal Reserve Board, October 1994.
5. Chen R.R., Scott L., "*Pricing Interest Rate Options in a Two-Factor Cox-Ingersoll-Ross Model of the Term Structure.*", The Review of Financial Studies, Vol. 5 1992, pp. 613-636.
6. Cheyette O., "*Term Structure Dynamics and Mortgage Valuation*", Journal of Fixed Income, March 1992.
7. Cox J.C., Ingersoll J.E., Ross S.A., "*A Theory of the Term Structure of Interest Rates.*", Econometrica, March 1985, pp. 385-407.
8. Duffie D., Kan R., "*A Yield-Factor Model of Interest Rates.*", Mathematical Finance, 1996, Vol. 6, No. 4, pp 379-406.

9. El Karoui N., Geman H., Lacoste V., "*On the Role of State Variables in Interest Rate Models*", Working Paper, May 1995.

10. Fong H.G., Vasicek O.A., "*Interest rate Volatility as a Stochastic Factor.*", Working Paper, Gifford Fong Associates, February 1991.

11. Heath D., Jarrow R., Morton A., "*Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation*", Econometrica, Vol. 60, 1992, pp. 77-105.

12. Langetieg T.C., "*A Multivariate Model of the Term Structure.*", Journal of Finance, March 1980, pp. 71-97.

13. Longstaff F.A., Schwartz E.D., "*Interest-Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model.*", Journal of Finance, September 1992, pp. 1259-1282.

14. Press, Teukolsky, Vetterlin, Flannery, "*Numerical Recipes in C*", Second Edition, Cambridge University Press.

15. Ritchken P., Sankarasubramanian L., "*Volatility Structures of Forward Rates and the Dynamics of the Term Structure*", Mathematical Finance, Vol. 5 No. 1.

16. Vasicek O.A., "*An Equilibrium Characterisation of the Term Structure.*", Journal of Financial Economics, Vol. 5 1977, pp. 177-188.

6. IMPLEMENTING THE HEATH JARROW AND MORTON APPROACH TO TERM STRUCTURE MODELLING[1]

6.1 INTRODUCTION

We have examined in previous chapters some of the unsatisfactory aspects of the short rate models and Duffie and Kan Affine Yield Models when they have to be calibrated accurately to options prices. The remaining chapters explore an alternative approach pioneered by Heath, Jarrow and Morton (1992), (HJM). As before, we distinguish between implementation and calibration. Implementation is the process of making a model produce prices and calibration is the process of choosing parameters such that the implementation produces model prices consistent with the market. Calibration cannot be achieved without efficient implementation. In this chapter we examine implementation of HJM models to fulfil two aims.

We will review HJM implementations to show that they are in general very complex. It is important to be aware of the implementation problems because even if tricks can be used to allow efficient calibration to simple options, the calibrated HJM model may have to be used in applications involving non-standard options where the tricks cannot be used. For example, we may calibrate a HJM model using approximate pricing formulae and proceed to price more complex options using numerical algorithms because the approximations either do not perform well or do not exist for more complex instruments. We review two main numerical techniques. We

---

[1] This chapter draws on joint work with Andrew Carverhill: Carverhill A. and Pang K., "*Efficient and Flexible Bond Option Valuation in the Heath, Jarrow and Morton Framework*", Journal of Fixed Income, September 1995, Vol. 5, No 2, pp 70-77.

review the tree construction and Monte Carlo Simulation techniques. We also introduce a simple and efficient technique, Martingale Variance Reduction, for increasing the accuracy of the Monte Carlo simulation prices for a given number of simulations.

We will illustrate that HJM implementation difficulties are the major drawback with the HJM approach. This forces practitioners to restrict their use of HJM models to sub-classes of the general approach that allow either analytical, or accurate approximate, formulae for options prices or efficient numerical methods.

Section 6.2 examines pricing using trees and Section 6.3 examines pricing using Monte Carlo simulations. The two sections show that both trees and Monte Carlo simulation techniques would be unsuitable to practitioners for calibration when the HJM approach is used in its full generality. Section 6.4 examines the sub-classes that allow more efficient methods. It examines Gaussian HJM models and a class we call Markovian HJM that allows the short rate process to become Markovian when the short rate is supplemented with additional state variables. Section 6.5 summarises.

6.2 HEATH, JARROW AND MORTON FORWARD RATE TREES

HJM models are characterised by the volatility factors, $\underline{\sigma}(t, T, \omega)$, that give the sensitivity of the forward rate term structure $\{f(t, T): T \geq t\}$ of unit changes in the Brownian motions that drive uncertainty in the model. The ω in $\underline{\sigma}(t, T, \omega)$ denotes the possible dependence of the volatility factors on the path taken by the Brownian motions. HJM show the short rate process is given by

$$dr(t) = \left\{ \mathcal{F}_t(0,t) + \int_0^t \left[\underline{\sigma}_t(u,t) \int_u^t \underline{\sigma}(u,v)dv + \underline{\sigma}(u,t)^2 \right] du + \left[\int_0^t \underline{\sigma}_t(u,t) d\underline{\widetilde{W}}(u) \right] \right\} dt + \underline{\sigma}(t,t) d\underline{\widetilde{W}}(t)$$

where the dependence on ω has been suppressed. The third term in the drift of the short rate process shows the drift in general depends on the path of the Brownian motions even when the volatility factors are deterministic. This presents a central difficulty when working within the HJM framework since it means the short rate process is in general non-Markovian and that it is not possible to construct a recombining short rate tree like those considered in Chapter 4. Thus without restricting to the very special cases where the volatility factors have been chosen to provide Markovian short rate process, numerical techniques based on the short rate process will be very inefficient.

Heath, Jarrow and Morton (1991), (HJM91), show how we can construct trees for the entire forward rate curve when there are one or two factors. HJM91 use their technique to calibrate various one-factor models in Heath, Jarrow and Morton (1994) which we have already reviewed in Chapter 3.

We will show that the HJM91 forward rate trees will, in general, be unsuitable for practical applications. This is because the HJM91 forward curve tree will in general be non-recombining. We show why the tree does not, in general, recombine even in the one factor case. Two factor forward rate trees are more complicated and in general also non-recombining.

For an one factor model, HJM91 show a forward rate tree for the risk-neutral measure can be constructed using the branching given by

$$
\begin{array}{ll}
 & f(t,T)+\sigma(t,T,f(t,T))\sqrt{\Delta t}+\delta(t,T)\Delta t \\
f(t,T) & \\
 & f(t,T)-\sigma(t,T,f(t,T))\sqrt{\Delta t}+\delta(t,T)\Delta t
\end{array}
$$

where

$$
\delta(t,T)\Delta t=\frac{\partial}{\partial T}\ln\left(\cosh\left(\int_{t+\Delta t}^{T}\sigma(t,u,f(t,u))\sqrt{\Delta t}\,du\right)\right) \tag{6.1}
$$

with risk-neutral probability $q_t(\Delta t)=\frac{1}{2}+O(\Delta t)$ that the up-branch will be taken and risk-neutral probability 1- $q_t(\Delta t)$ that the down-branch will be taken. The difference between the instantaneous rates, starting at $f(t, T)$, after a up-down and a down-up stepping sequence is

$$
2\left[\sigma(t,T,f(t,T))-\sigma(t+\Delta t,T,f(t+\Delta t,T))\right]\sqrt{\Delta t}\,.
$$

Thus the instantaneous forward rate tree does not recombine except for the case when $\sigma(t, T, f(t, T)) \equiv \sigma$ which corresponds to the continuous time limit of the Ho-Lee (1986). Note that this branching structure will not recombine for the exponential Vasicek volatility structure given by $\sigma(t,T)=\sigma\exp(-\alpha(T-t))$. This is so even though it is possible to construct a recombining short rate tree (and hence a recombining forward rate tree since the short rate is the only state variable) for the Extended Vasicek short rate model. There is no contradiction since the extended Vasicek recombining short rate tree is trinomial and has time and state dependent branching probabilities and branching structure that are specially chosen to ensure the tree recombines. The HJM91 tree has been designed to work for general volatility structures.

To see why the HJM91 tree is impractical, observe that their tree will have $2^{n+1}-1$ nodes after the nth time step. The number of tree nodes grow

exponentially. This produces severe consequences for pricing and calibration since, even after the tree has already been constructed, pricing just one European option will require an enormous amount of calculation to evaluate the expected discounted terminal value of the option.

We have devised some numerical experiments to investigate the practicalities of their approach. In order to be able to assess the accuracy unambiguously, we shall consider an example which permits analytic solutions.

We will use the HJM91 tree to price options on PDBs when the forward rate volatilities are given by the exponential Vasicek volatility $\sigma(t, T) = \sigma \exp[-\alpha(T-t)]$. We make this choice of volatility to allow us to compare prices obtained using the forward rate tree with the analytical solutions provided by Jamshidian (1987) and Hull and White (1990). Those papers show the time t price of a T-maturity European Call option on a s-maturing PDB for this choice of volatility function is given by

$$P(t,s)N(h) - XP(t,T)N(h-\sigma_P)$$

where

$$h = \frac{1}{\sigma_P}\ln\frac{P(t,s)}{XP(t,T)} + \frac{\sigma_P}{2}$$

$$\sigma_P = \sqrt{\frac{\sigma^2\left(1-e^{-\alpha(s-T)}\right)^2\left(1-e^{-2(T-t)}\right)}{2\alpha^3}}$$

and X is the strike price. The value of the corresponding European Put option is given by

$$XP(t,T)N(-h+\sigma_P) - P(t,s)N(-h).$$

Of course, for this simple valuation problem, it would be unnecessary to construct the forward rate tree, but the exercise illustrates the difficulties

encountered if one were to attempt to calibrate the forward rate volatilities using forward rate trees.

Table 1 shows the time taken for different size trees to price a variety of one year options on an 8 year PDB with strike equal to the one year forward price of the 8 year PDB. It shows a number of interesting results:

1) The time taken to price the options grows extremely rapidly and that for a sixteen step tree, the time taken is about sixty-eight minutes. The reported computation times are those taken on a UNIX terminal to price all four options. Note that these reported times are lower than what would be achieved in practice: Whereas for the Vasicek volatility function we can evaluate equation 6.1 analytically, for a more general functional form for the volatility function, we would have to integrate equation 6.1 numerically and increase the computational effort needed to construct the tree.
2) A small forward rate tree can price European options on PDBs accurately. This is because even with a few steps, the non-recombining tree provides many samples for the option value. For example, a 10 step tree returns the price of the option as the average of $2^{10} = 1024$ sample values.
3) The forward rate tree does not price American options very well. The variation in prices as the tree increases from ten and sixteen steps trees are small. However, for five out of the seven trees constructed, the trees give American call prices that are smaller than the exact theoretical minimum; the exact price of an European call. Thus the American call options are priced poorly. Forward rate trees are not able to price the American option so well because they do not allow for many early exercise opportunities. For example, a ten step tree only allows for nine early exercise opportunities.

4) Table 1 suggests that it may be possible to use a small forward rate tree, about ten steps in size, to calibrate HJM models to options prices. Indeed Amin and Morton (1994) use forward rate trees to calibrate to short maturity American options on Eurodollar futures to extract implied volatility term structures.

Note, however, that most interest rate derivatives are sensitive to the correlation structure of zero-rate changes but a single-factor model implies that all zero-rate changes are perfectly correlated. Single factor models are inappropriate except for the simplest options. Thus, in practice, practitioners would need to calibrate at least a two factor model. HJM (1991) also show we can approximate discretely a two-factor model by constructing a trinomial non-recombining tree that, after n steps, would have $(3^{n+1}-1)/2$ tree nodes. A trinomial tree, after ten steps, has 88573 tree nodes whereas the binomial tree has only 2047. Thus pricing and calibrating with a two-factor forward rate tree would require far more time than a one-factor tree.

Our example shows that a forward rate tree is generally impractical and unsuitable to practitioners for calibration since a calibration typically involves an optimisation of some parameters to minimise a weighted residual sum of squares between market and model prices. The optimisation would require the construction of a new tree for every candidate volatility structure and depending on the chosen routine and the problem, the optimisation may involve many iterations before arriving at the optimal parameters.

6.3 MONTE CARLO SIMULATIONS

6.3.1 Simple Monte Carlo Simulations

The method of Monte Carlo simulations, in its most basic, entails simulating the path of the underlying variables determining the option value to the maturity option. For example, consider a derivative that at maturity has value $F(T)$ that is determined by the path of the interest rate term structure up to time T. It's value is given by

$$F(0) = \tilde{E}\left[\exp\left\{-\int_0^T r(u)du\right\}F(T)\right].$$

Like before, we assume that $\underline{\sigma}(s, u, \omega) = \underline{\sigma}(s, u, f(s, u))$ so that the risk-neutral process for the forward rates are given by

$$df(t,T) = \alpha(t,T,\cdot)dt + \underline{\sigma}(t,T,f(t,T)) \cdot d\underline{\tilde{W}}(t) \tag{6.2}$$

where

$$\alpha(t,T,\cdot) = \underline{\sigma}(t,T,f(t,T)) \cdot \int_t^T \underline{\sigma}(t,v,f(t,v))dv. \tag{6.3}$$

Monte Carlo simulations typically proceed as follows. We generate sample paths for the instantaneous forward rates at discrete intervals by dividing the time to maturity T into n equal segments so that the length of time step Δt is T/n. Suppose we only need forward rates up to a maximum maturity $M\Delta t$. Then we can simulate the evolution of the forward rate curve using

$$f((k+1)\Delta t, j\Delta t) = f(k\Delta t, j\Delta t) + \alpha(k\Delta t, j\Delta t, \cdot)\Delta t + \underline{\sigma}(k\Delta t, j\Delta t, f(k\Delta t, j\Delta t))\sqrt{\Delta t}\,\underline{\varepsilon}(k)$$

$$\alpha(k\Delta t, j\Delta t, \cdot) = \underline{\sigma}(k\Delta t, j\Delta t, f(k\Delta t, j\Delta t))\sum_{m=k}^{j-1}\underline{\sigma}(k\Delta t, m\Delta t, f(k\Delta t, m\Delta t))\Delta t$$

where for each $k = 0, \ldots, n\text{-}1$, $j = k+1, \ldots, M$. $\underline{\varepsilon}(k)$ is the kth sample vector of multivariate normal random variables with distribution $N(0, \mathrm{I})$.

Note that the short rate at time $k\Delta t$, $r(k\Delta t) \equiv f(k\Delta t, k\Delta t)$, is assumed to apply for the interval $[k\Delta t, (k+1)\Delta t)$. Similarly the forward rate $f(k\Delta t, j\Delta t)$ is assumed to apply for the interval $[j\Delta t, (j+1)\Delta t)$. Thus the time $k\Delta T$ value of a $m\Delta t$ maturity bond is given by

$$P(k\Delta t, m\Delta t) = \exp\left\{-\Delta t \sum_{i=k}^{m-1} f(k\Delta t, i\Delta t)\right\}.$$

In this way, we can evaluate bond prices, zero coupon yields or any other function of the term structure. Thus given a sample path, we can evaluate the option payoff and discount back to the present for any simulated path. The current value of the option is given by the average of many sample values obtained from many sample paths.

The accuracy of the simulated price depends on the total number of simulations or samples for the discounted payoff. The *standard error of the* sample mean is $O(1/\sqrt{N})$ when the option value is obtained by average N sample values. Very often in application, a very large number of simulations are needed to reduce the standard error to acceptable levels. Thus when time is limited, it is important that the simulations be made efficient.

6.3.2 Efficient Monte Carlo Simulations

The simple technique just presented is very unlikely to be used in practice. This is because simulating the entire forward rate curve is very often more than what is required for pricing interest rate derivatives. The value of interest rate derivatives is usually determined by the value of zero coupon yields or bond prices that are functions of averages of the instantaneous forward rates. Thus it is more efficient to simulate the few bond prices that determine the option value than to simulate the entire forward rate curve. We will examine cases when the bond returns

volatilities are stochastic. We consider the special case when volatilities are deterministic in Section 6.4.

Monte Carlo simulations are inherently time-consuming and so in order to speed the simulations up, it may be necessary to simplify the model. We shall consider some simplification that will enable faster simulations. Consider, first the case if we were to assume the forward rate process given by equations 6.2 and 6.3. Then the bond returns process would be given by

$$\frac{dP(t,T)}{P(t,T)} = rdt + \underline{v}(t,T) \cdot d\underline{\widetilde{W}}(t) \tag{6.4}$$

with

$$\underline{v}(t,T\cdot) = -\int_t^T \underline{\sigma}(t,s,f(t,s))ds\,. \tag{6.5}$$

Equation 6.5 shows that not only are the bond returns volatility factors computationally intensive to evaluate, but also they require the entire forward rate curve that would have to simulated too. This defeats the whole purpose of working with the bond prices. Therefore, practitioners tend to obtain stochastic volatility models differently.

The simplifications we shall use are designed to enable more efficient Monte Carlo simulations. Thus to illustrate the motivation for our simplifications, we first consider how we plan to conduct the Monte Carlo simulations. We improve the accuracy of our simulations using a technique called Martingale Variance Reduction. Clewlow and Carverhill (1994) provide an early example of their use in financial applications. The techniques differ by their choice of martingales. The stochastic volatility examples are provided in Sections 6.3.2.2 and 6.3.2.3.

6.3.2.1 Martingale Variance Reduction

This section follows closely Carverhill and Pang (1995). We work with a forward risk-adjusted measure here whereas Carverhill and Pang (1995) work with the risk-neutral measure.

Consider the valuation of European options on coupon bonds. This is an important example since, as we show in chapter 7, swaptions are equivalent to options on coupon bonds.

Let the European call and put options have maturity T and let the underlying coupon bond pay cash amounts c_i at times q_i, $i = 1, \ldots, n$, with $T < q_1 < q_2 \ldots < q_{n-1} < q_n = q$. Define the money account $\beta(t)$ by $\beta(t) = \exp\left\{\int_0^t r(s)ds\right\}$. Then the time t prices of the European options are given by

$$p(t) = \widetilde{E}_t\left[\frac{\beta(t)}{\beta(T)}\left(I_f\left\{\sum_{i=1}^{n} c_i P(T,q_i) - X\right\}\right)^+\right]$$

where I_f takes value 1 for a call and -1 for a put and $P(T, q_i)$ is the time T solution to

$$\frac{dP(t,q_i)}{P(t,q_i)} = rdt + \underline{v}(t,q_i)\cdot d\underline{\widetilde{W}}(t)\,. \tag{6.6}$$

We see that the risk neutral bond price process involves the short rate which is in general path dependent and requires the whole of the initial instantaneous forward rate curve to determine its evolution. The presence of the short rate makes simulation awkward and suggests a change of numéraire. The natural choice is the T maturity PDB so that the prices of the options are given by

$$p(t) = P(t,T)E_t^T\left[\left(I_f\left\{\sum_{i=1}^{n} c_i P(T,q_i) - X\right\}\right)^+\right]$$

where E_t^T denotes expectation with respect to the equivalent martingale measure induced by taking the T-maturity PDB as the numéraire. El Karoui et al (1995) show that the risk-neutral measure and the T-measure are related by the Radon-Nikodym derivative

$$\frac{d\underline{Q}^T}{d\tilde{Q}} = \frac{P(T,T)/P(0,T)}{\beta(T)/\beta(0)}$$

$$= \exp\left\{\int_0^T \underline{v}(u,T)d\underline{\tilde{W}}(u) - \frac{1}{2}\int_0^T \left|\underline{v}(u,T)\right|^2 du\right\}$$

so that by Girsanov's Theorem

$$d\underline{W}^T(t) = d\underline{\tilde{W}}(t) - \underline{v}(t,T)dt \qquad \textbf{(6.7)}$$

is a Brownian motion with respect to the T-measure. Substituting equation 6.7 into equation 6.6 gives

$$P(T,q_i) = \frac{P(t,q_i)}{P(t,T)}\exp\left\{\int_t^T \left[\underline{v}(u,q_i) - \underline{v}(u,T)\right]d\underline{W}^T(u) - \frac{1}{2}\int_t^T \left|\underline{v}(u,q_i) - \underline{v}(u,T)\right|^2 du\right\}$$

so that the values of the European call and put options are given by

$$p(t) = P(t,T)E_t^T\left[\left(I_f\left\{\sum_{i=1}^{n} c_i \frac{P(t,q_i)}{P(t,T)} M_t^{T,q_i} - X\right\}\right)^+\right] \qquad \textbf{(6.8)}$$

where

$$M_t^{T,q_i} = \exp\left\{\int_t^T \left[\underline{v}(u,q_i) - \underline{v}(u,T)\right]d\underline{W}^T(u) - \frac{1}{2}\int_t^T \left|\underline{v}(u,q_i) - \underline{v}(u,T)\right|^2 du\right\}. \qquad \textbf{(6.9)}$$

If we were to simulate the martingales defined by equation 6.9 then equations 6.8 and 6.9 would enable us to obtain a simple Monte Carlo

estimate of the value of the options. Instead consider the control variate estimate $p_j^{cv}(0)$ defined by

$$p_j^{cv}(0) = p_j(0) + \sum_{i=1}^{n} \beta_i \left(M_0^{T,q_i}(j) - 1 \right) \tag{6.10}$$

where $M_0^{T,q_i}(j)$ is the realised value of the martingale M_0^{T,q_i} on the jth sample path, $p_j(0)$ is jth sample value defined by equations 6.8 and 6.9 and β_i are constants for $i = 1, \ldots, n$. Note that

$$E^T[p_j^{cv}(0)] = E^T[p_j(0)]$$

since $E^T[M_0^{T,q_i}(j)] = 1$ for all i and that

$$Var^T[p_j^{cv}(0)] = Var^T[p_j(0)] + \sum_{i=1}^{n} \beta_i Cov^T[p_j(0), M_0^{T,q_i}(j)]$$

$$+ \sum_{i=1}^{n} \sum_{k=1}^{n} \beta_i \beta_k Cov^T[M_0^{T,q_i}(j), M_0^{T,q_k}(j)].$$

$p_j^{cv}(0)$ is the control variate estimate of $p_j(0)$. $p_j^{cv}(0)$ is an unbiased estimate of the option value and its standard error can be minimised by choosing the β's appropriately. It is clear that we can choose $\underline{\beta}$ such that se($p_j^{cv}(0)$) ≤ se($p_j(0)$) since the equality is obtained when $\underline{\beta} = \underline{0}$. We can estimate $\underline{\beta}$ efficiently by proceeding as follows. With implicit sign changes, we can write

$$p_j(0) = E_0^T[p_j^{cv}(0)] + \sum_{i=1}^{n} \beta_i \left[M_0^{T,q_i}(j) - 1 \right] + \varepsilon_j$$

where ε_j are identically and independently distributed noise terms with mean 0. Performing an ordinary least squares regression, (OLS), of $p_j(0)$ on an intercept and $\left[M_0^{T,q_i}(j) - 1 \right]$, $i = 1, \ldots, n$, gives $E_0^T[p_j^{cv}(0)]$ as the intercept and the se($p_j^{cv}(0)$) as the standard deviation of the regression intercept. This technique of adding martingale control variates onto the your variable

to reduce its standard error is called Martingale Variance Reduction, (MVR). The more correlated the martingale control variates are with your variable, the more effective the reduction in the standard error.

We now return to the stochastic volatility examples. Following our discussion on the MVR technique, it follows that the technique would work best if we can simulate the control variates, $\left\{M_0^{T,q_i}(j)-1\right\}$, $i = 1, \ldots, n$, easily. Equation 6.9 shows that this depend on the choice of the bond return volatility factors. The bond return volatility factors can be simulated most easily if they depended only on a small number of Markovian variables. Our examples have this feature. We use the Fong and Vasicek (1991), (FV), and Longstaff and Schwartz (1992), (LS), models for our examples. We compare the simulations results between a simple Monte Carlo Simulation and a Monte Carlo Simulation with MVR. Note, however, that our methods can be applied to more general models but we perform the simulations for the FV and LS models because they are familiar.

6.3.2.2 Stochastic Volatility Example 1: Fong and Vasicek (1991)

Consider the HJM model defined by

$$\frac{dP(t,q)}{P(t,q)} = r(t)dt + \sigma(t)\sum_{j=1}^{n}\kappa_j(t,q)d\widetilde{W}_j(t) \qquad \textbf{(6.11)}$$

with

$$d\sigma(t) = \xi(\sigma(t))dt + \eta(\sigma(t))d\widetilde{W}_0(t)$$

and

$$d\widetilde{W}_0(t)d\widetilde{W}_j(t) = \rho_j dt\,.$$

$\kappa_j(t, q)$, $j = 1, \ldots, n$, are deterministic functions that when scaled by $\sigma(t)$ give the volatility factors. The FV model falls into this category. FV, in effect, assume the following pair of risk-neutral processes

$$dr(t) = [\alpha(\bar{r} - r(t)) + \lambda v(t)]dt + \sqrt{v(t)}d\tilde{B}_1(t)$$

$$dv(t) = (\gamma + \xi\eta)\left[\frac{\gamma}{\gamma + \xi\eta}\bar{v} - v(t)\right]dt + \xi\sqrt{v(t)}d\tilde{B}_2(t)$$

where $d\tilde{B}_1(t)$ and $d\tilde{B}_2(t)$ are Brownian increments with respect to the risk-neutral measure and $d\tilde{B}_1(t)d\tilde{B}_2(t) = \rho dt$. FV are able to solve for a bond pricing formula from which it is easy to obtain the bond return process

$$\frac{dP(t,q)}{P(t,q)} = r(t)dt - \sqrt{v(t)}\begin{bmatrix} D(q-t) \\ F(q-t)\xi \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix} d\underline{\tilde{W}}(t)$$

where $D(\cdot)$ and $F(\cdot)$ are deterministic functions and $d\underline{\tilde{W}}(t)$ are independent Brownian increments with respect to the risk-neutral measure. Thus setting $\sigma(t) = \sqrt{v(t)}$ and

$$\underline{\kappa}(t,q) = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix}^T \begin{bmatrix} D(q-t) \\ F(q-t)\xi \end{bmatrix}$$

in equation 6.11 makes the resulting HJM model equivalent to FV. The HJM model of equation 6.11 of course allow for a much wider class of models than just FV.

Table 2 shows the simulations results for the FV model. Table 2 uses the FV parameters $\alpha = 1.5$, $\xi = 0.1$, $\rho = 0.5$, $r_0 = 0.1$, $\bar{r} = 0.1$, $v_0 = 0.01$, $\bar{v} = 0.01$, $\gamma = 1.0$ and $\lambda = 0$. From Table 2 we see that the variance reduction works better for the coupon bond than for pure discount bond; the reduction in the standard errors for the options on a coupon bond is by a factor of four compared with two for the options on the PDB. These reductions indicate savings by factors of about sixteen and four,

respectively, in the number of iterations required to achieve a given accuracy. The reason that the variance reduction is more effective for the coupon bond is simply that when there are coupons there are more control variates. Recall, however, that using these variates incurs hardly any extra cost, because they have already been calculated in the Monte Carlo procedure itself. The only extra overhead is the calculation of an OLS regression. On an UNIX terminal, 50000 simple Monte Carlo simulations using 20 steps per year took 4200 and 4350 seconds to price the options on the PDB and coupon bond respectively; the additional OLS regression used to generate the control variate estimates took an additional 50 and 290 seconds for the options on the PDB and coupon bond respectively. The figures show the MVR technique is able to reduce the standard errors with very little extra work.

6.3.2.3 Stochastic Volatility Example 2: Longstaff and Schwartz (1992)

We can extend the HJM model of equation 6.11 by defining the bond return process

$$\frac{dP(t,q)}{P(t,q)} = r(t)dt + \sum_{i=1}^{n} \sigma_i(t)\kappa_i(t,q)d\widetilde{W}_i(t) \qquad \textbf{(6.12)}$$

where $\underline{\sigma}(t, \underline{\widetilde{W}}(t))$ follows a well-defined Markovian process. Longstaff and Schwartz (1992), LS, fall into this category. Essentially, LS model an economy in which $r(t)$ and $v(t)$, the volatility of the short rate, are related to the variables $x(t)$ and $y(t)$ by $r(t) = \alpha x(t) + \beta y(t)$ and $v(t) = \alpha^2 x(t) + \beta^2 y(t)$ where $x(t)$ and $y(t)$ follow the risk-neutral processes

$$dx(t) = (\gamma - \delta x(t))dt + \sqrt{x(t)}d\widetilde{W}_1(t)$$

$$dy(t) = (\eta - \varepsilon x(t))dt + \sqrt{y(t)}d\widetilde{W}_2(t).$$

LS obtain a bond pricing formula from which it is easy to show that bond return follows the process

$$\frac{dP(t,q)}{P(t,q)} = r(t)dt + \begin{bmatrix} \alpha\sqrt{x(t)}\big(C(q-t)+\alpha D(q-t)\big) \\ \beta\sqrt{y(t)}\big(C(q-t)+\beta D(q-t)\big) \end{bmatrix}^T d\underline{\widetilde{W}}(t)$$

where $C(\cdot)$ and $D(\cdot)$ are deterministic functions and so the formulation for equation 6.12 follows immediately.

Table 3 shows results for the LS model with parameter values α = 0.05, β = 0.06, γ = 0.8, δ = 0.8, ε = 0.5, η = 0.9, r_0 = 0.1 and υ_0 = 0.055. The results for LS are essentially the same as for the FV implementation. The variance reduction is better for the coupon bond, and requires fewer iterations for a given accuracy, by a factor of about nine.

We have illustrated the technique of Monte Carlo simulations by pricing European call options on a PDB and on a coupon bond. The technique applies equally easily to all other European options including the path-dependent variety. Monte Carlo techniques cannot however be applied easily for the pricing of American options because it is necessary to determine the early exercise boundary. Some researchers have been tackling the problem of pricing of American options with some success. See for example, Broadie and Glasserman (1994). However, Monte Carlo simulation techniques cannot in practice be used for calibration. Firstly, the technique takes a long time to produce accurate prices and secondly, many optimisation routines are not robust enough to handle function values with noise. For example, a gradient based optimisation routine may not converge because the small errors in prices can produce large errors in gradients.

Monte Carlo techniques are probably best reserved for pricing complex instruments after the model has been suitably calibrated.

6.4 HJM SIMPLIFICATIONS

We have examined pricing in the HJM approach using forward rate trees and Monte Carlo simulations in Sections 6.2 and 6.3 respectively. Although those pricing algorithms are general and suitable for pricing a wide variety of interest rate derivatives, they are certainly unsuitable for calibration work since they are far too slow. In practice, it is necessary to find simplifications or alternative methods to allow faster methods. In this section, we examine two types of simplifications that allow much faster pricing. The simplifications amount to restricting the type of volatility functions we permit.

We examine Gaussian HJM models and simplifications provided by Cheyette (1991) and Ritchken and Sankarasubramanian (1995). Both Cheyette (1991) and Ritchken and Sankarasubramanian (1995) find supplementary variables that allow the short rate process to become jointly Markovian with respect to itself and supplementary variables. The supplementary variables allow for the construction of recombining trees and for the derivation of partial differential equations for the values of some derivatives that can be solved by standard finite difference methods.

6.4.1 Gaussian Heath, Jarrow and Morton and Numerical Integration

For Gaussian HJM models we can sometimes derive analytic solutions or accurate approximations. When we cannot we can often eliminate the need for Monte Carlo simulations and perform a more efficient numerical integration instead. Rather than using Monte Carlo simulations to perform what is in effect a numerical integration, we make use of the

analytic tractability provided by the Gaussian assumptions to determine various distributions exactly to allow a more efficient numerical integration.

Consider again the prices of European call and put options on the coupon bond of Section 6.3.2 that are given by

$$p(t) = P(t,T)E_t^T\left[\left(I_f\left\{\sum_{i=1}^{n} c_i \frac{P(t,q_i)}{P(t,T)} M_t^{T,q_i} - X\right\}\right)^+\right] \tag{6.8}$$

where

$$M_t^{T,q_i} = \exp\left\{\int_t^T \left[\underline{v}(u,q_i) - \underline{v}(u,T)\right] d\underline{W}^T(u) - \frac{1}{2}\int_t^T \left|\underline{v}(u,q_i) - \underline{v}(u,T)\right|^2 du\right\}. \tag{6.9}$$

We cannot price these options analytically in general even when interest rates are Gaussian although it is possible, as we show in Chapter 7, to derive good approximations when the coupons are small.

Examining equations 6.8 and 6.9 show, when the volatilities are deterministic, we do not have to simulate for we can price the European options using

$$p(0) = \frac{1}{N}\sum_{j=1}^{N} p_j(0) \tag{6.13}$$

where the jth sample value of the option is given by

$$p_j(0) = P(0,T)\left(I_f\left\{\sum_{i=1}^{n} c_i \frac{P(0,q_i)}{P(0,T)} e^{\underline{\varepsilon}_i(j)} - X\right\}\right)^+ \tag{6.14}$$

and $\underline{\varepsilon}(j)$ is the jth sample of a vector of multivariate normal random variables with distribution $N(\underline{\mu}, Q)$ where

$$\underline{\mu}_i = -\frac{1}{2}\int_t^T \left|\underline{v}(u,q_i) - \underline{v}(u,T)\right|^2 du \tag{6.15}$$

and

$$Q_{mn} = \int_0^T \left[\underline{v}(u,q_m) - \underline{v}(u,T)\right]\left[\underline{v}(u,q_n) - \underline{v}(u,T)\right]du. \qquad (6.16)$$

Equations 6.13 to 6.16 provide an estimate that has a standard error of $O(1/\sqrt{N})$. It is well known that a numerical integration of the type presented here can be made much more efficient by using low discrepancy sequences. For example, see Niederreiter (1988). By replacing the multivariate normal samples of equation 6.14 with a low discrepancy series, it is possible to obtain an estimate with an standard error of $O(1/N)$. Paskov (1994) and Joy, Boyle and Tan (1995) have applied low-discrepancy sequences to financial problems.

The numerical integration provided by equations 6.13 to 6.16 is far more efficient than a Monte Carlo simulation. The method presented, however, can only be used when a Gaussian HJM model is appropriate. When the Gaussian assumptions are unsuitable, we cannot usually avoid the more intensive simulations.

6.4.2 Markovian Heath, Jarrow and Morton

We have seen that a key contributing factor to the difficulty with HJM implementations is the non-Markovian behaviour of the short rate. Cheyette (1992) and Ritchken and Sankarasubramanian (1995) propose a similar technique to circumvent the problem. They identify classes of HJM models, characterised by their volatility functions, that when supplemented with additional state variables, allow the short rate to follow a joint Markov process. Importantly, volatilities can be stochastic. We call this type of models Markovian HJM.

For an one-factor HJM model, Cheyette (1992) shows for the class of HJM models where the volatility is separable into a time-dependent, a rate-dependent term and a maturity-dependent term

$$\sigma(t,T) = \frac{\sigma_r(t)}{f(t)} f(T) \tag{6.17}$$

where $\sigma_r(t)$ is the volatility of the short rate, it is possible to supplement r with an additional variable to give a pair of state variables for the model. This supplementary variable is given by

$$v^2(t) = f^2(t) \int_0^t \frac{\sigma_r^2(s)}{f^2(s)} ds .$$

Cheyette (1992) shows r and v follow the two-dimensional Markovian process given by

$$d[r(t) - f(0,t)] = \left[v^2(t) + \frac{d \ln f(t)}{dt} (r(t) - f(0,t)) \right] dt + \sigma_r(t) d\widetilde{W}(t)$$

$$dv^2(t) = \left[\sigma_r^2(t) + 2 \frac{d \ln f(t)}{dt} v^2(t) \right] dt .$$

It can also be shown that the price, $P_A(r, v^2, t)$ of a security, with cashflows $C_A(r, v^2, t)$ that depend only on the forward rate curve at time t, obeys, subject to the appropriate boundary conditions, the two-dimensional partial differential equation given by

$$\frac{\partial P_A}{\partial t} + \frac{\sigma_r^2}{2} \frac{\partial^2 P_A}{\partial r^2} + \left\{ \frac{df(0,t)}{dt} + v^2 - \frac{r - f(0,t)}{\tau(t)} \right\} \frac{\partial P_A}{\partial r}$$

$$+ \left\{ \sigma_r^2 - \frac{2v^2}{\tau(t)} \right\} \frac{\partial P_A}{\partial (v^2)} - rP_A + C_A(r, v^2, t) = 0$$

where

$$\frac{1}{\tau(t)} \equiv -\frac{d \ln f(t)}{dt} .$$

Thus the HJM model with volatility defined by equation 6.17 is amenable to efficient numerical methods. The above approach can be extended easily to multifactor HJM models. For every factor, the state space is supplemented with an additional variable so that a n-factor model will

have $2n$ state variables that follow a joint Markov process. Security prices are solutions to a $2n$ dimensional partial differential solution. Ritchken and Sankarasubramanian (1995) obtains similar results and shows that

$$\sigma(t,T) = \sigma_r(t)\exp\left[-\int_t^T \kappa(x)dx\right]$$

is a necessary and sufficient condition to allow prices to be completely determined by the two-state variable Markov process.

The Markovian HJM models form a significant contribution to the range of models practitioners can use as they are far easier to calibrate. American options can be priced using standard finite difference methods to solve the partial differential equation and Monte Carlo simulations can be used easily to price path-dependent options. However, calibration is still time-consuming. There will be very few analytical prices for options except for the Gaussian cases. Practitioners will probably want at least two-factors to capture the correlation structure of zero coupon yield changes and the calibration may therefore involve solving 4-dimensional partial differential equations.

The Markovian HJM models do carry their own disadvantages. The restriction on the volatility function may preclude certain desirable forms. This is particularly true if we need volatility factors that have stochastic levels but time stationary shapes.

6.5 SUMMARY

We have examined HJM implementations in this chapter to illustrate the computational challenges involved in pricing interest rate derivatives. We have shown that both forward rate trees and Monte Carlo simulations are too slow to be practical for calibrating HJM models in general. This

observation forces practitioners to use certain subclasses of the HJM approach that provide either analytic tractability or computational efficiency. For analytic tractability, we examined the Gaussian HJM models and for computational efficiency, we examined the Markovian HJM models.

6.6 REFERENCES

1. Amin K.I. and Morton A.J., "*Implied Volatility Functions in Arbitrage-Free Term Structure Models*", Journal of Financial Economics, Vol. 35, 1994.
2. Broadie M. and Glasserman P., "*Pricing American-Style Securities Using Simulation*", Working Paper, Columbia University, 1994.
3. Carverhill A. and Pang K., "*Efficient and Flexible Bond Option Valuation in the Heath, Jarrow and Morton Framework*", Journal of Fixed Income, September 1995, Vol. 5, No 2, pp 70-77.
4. Cheyette O., "*Term Structure Dynamics and Mortgage Valuation*", Journal of Fixed Income, March 1992.
5. Clewlow L.C. and Carverhill A., "*On the Simulation of Contingent Claims*", Journal of Derivatives, Winter 1994, pp. 66-74.
6. Fong H.G. and Vasicek O., "*Interest Rate Volatility as a Stochastic Factor*", Working Paper, Gifford Fong Associates, 1991.
7. Heath D., Jarrow R., and Morton A.., "*Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation*", Econometrica, Vol. 60, No. 1, 1992.
8. Heath D., Jarrow R., and Morton A.., "*Contingent Claim Valuation with a Random Evolution of Interest Rates*", Journal of Futures Markets, Vol. 10, 1991.

9. Ho T.S.Y. and Lee S.B., "*Term Structure Movements and Pricing Interest Rate Contingent Claims*", Journal of Finance, 1986, Vol. XLI, No. 5, pp 1011-1029.

10. Hull J. and White A., "*Pricing Interest Rate Derivative Securities*", The Review of Financial Studies, Vol. 3 (4), 1990.

11. Jamshidian, F., "*An Exact Bond Option Formula*", Journal of Finance, March 1989, Vol. XLIV, No.1 , pp 205-209.

12. Joy C., Boyle P. and Tan K.S., "*Quasi Monte Carlo Methods in Numerical Finance*", Working paper, February 1995, University of Waterloo.

13. Longstaff F.A., Schwartz E.D., "*Interest-Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model.*", Journal of Finance, September 1992, pp. 1259-1282.

14. Niederreiter H., "*Numerical Integration III*", ISNM vol. 85, Bräss H. and Hämmerlin G., eds. (Basel: Birkhaüser), pp. 157-171.

15. Paskov S.H., "*Computing High Dimensional Integrals with Applications to Finance*", Technical Report CUCS-023-94, October 1994, Columbia University.

16. Ritchken P. and Sankarasubramanian L., "*Volatility Structures of Forward Rates and the Dynamics of the Term Structure*", Mathematical Finance, Vol. 5 No. 1.

	Table 1: Pricing with a HJM Forward Rate Tree						
	1 Year Options an 8 Year Pure Discount Bond in Vasicek Framework, (Exact European price 0.006851)						
Steps / year	**10**	**11**	**12**	**13**	**14**	**15**	**16**
Time Taken, sec	3	6	15	47	77	1543	4060
European Call	0.006810	0.06975	0.006824	0.006952	0.006833	0.006934	0.006840
European Put	0.006810	0.06975	0.006824	0.006952	0.006833	0.006934	0.006840
American Call	0.006819	0.006983	0.006832	0.006958	0.006840	0.006941	0.006845
American Put	0.047463	0.047462	0.047461	0.047461	0.047460	0.047460	0.047459

Table 2: Fong & Vasicek Monte Carlo Simulation Prices Three Year European Call on Six Year Pure Discount Bond							
20 steps/year, 50000 simulations				100 steps/year, 50000 simulations			
Strike	0.738147	0.745603	0.753059	**Strike**	0.738147	0.745603	0.753059
Moneyness	99%	100%	101%	**Moneyness**	99%	100%	101%
Simple MC	0.010815	0.007504	0.004889	**Simple MC**	0.010872	0.007599	0.004998
Std Error	0.000058	0.000049	0.000040	**Std Error**	0.000059	0.000050	0.000042
MC with MVR	0.010766	0.007463	0.004855	**MC with MVR**	0.010953	0.007665	0.005050
Std Error	0.000027	0.000027	0.000026	**Std Error**	0.000028	0.000028	0.000027
Three Year European Call on Six Year 10% Semi-Annual Coupon Bond							
20 steps/year, 50000 simulations				100 steps/year, 50000 simulations			
Strike	0.989276	0.999269	1.009262	**Strike**	0.989276	0.999269	1.009262
Moneyness	99%	100%	101%	**Moneyness**	99%	100%	101%
Simple MC	0.014142	0.009687	0.006197	**Simple MC**	0.014215	0.009809	0.006338
Std Error	0.000075	0.000063	0.000052	**Std Error**	0.000076	0.000065	0.000053
MC with MVR	0.014090	0.009644	0.006164	**MC with MVR**	0.014359	0.009933	0.006439
Std Error	0.000014	0.000016	0.000017	**Std Error**	0.000015	0.000016	0.000017

Table 3: Longstaff & Schwartz Monte Carlo Simulation Prices							
Three Year European Call on Six Year Pure Discount Bond							
20 steps/year, 50000 simulations				100 steps/year, 50000 simulations			
Strike	0.646075	0.652601	0.659127	Strike	0.646075	0.652601	0.659127
Moneyness	99%	100%	101%	Moneyness	99%	100%	101%
Simple MC	0.020339	0.017744	0.015321	Simple MC	0.020581	0.017977	0.015545
Std Error	0.000113	0.000104	0.000096	Std Error	0.000113	0.000105	0.000097
MC with MVR	0.020390	0.017791	0.015363	MC with MVR	0.020552	0.017951	0.015522
Std Error	0.000057	0.000056	0.000055	Std Error	0.000057	0.000056	0.000055
Three Year European Call on Six Year 10% Semi-Annual Coupon Bond							
20 steps/year, 50000 simulations				100 steps/year, 50000 simulations			
Strike	0.879675	0.888560	0.897446	Strike	0.879675	0.888560	0.897446
Moneyness	99%	100%	101%	Moneyness	99%	100%	101%
Simple MC	0.025770	0.022224	0.018932	Simple MC	0.026076	0.025518	0.019213
Std Error	0.000142	0.000131	0.000120	Std Error	0.000143	0.000132	0.000120
MC with MVR	0.025858	0.022305	0.019005	MC with MVR	0.026056	0.025501	0.019199
Std Error	0.000046	0.000047	0.000047	Std Error	0.000047	0.000048	0.000048

7. CALIBRATING GAUSSIAN HEATH JARROW AND MORTON

7.1 INTRODUCTION

Our examination, in Chapter 6, of a selection of numerical methods for pricing interest rate derivatives in the HJM framework shows that unless one considers special subclasses of the HJM approach, that allow closed-form solutions or efficient numerical methods, then one would not be able to use HJM in practice because the calibration would be too slow. We considered two subclasses that allow easier calibration: Gaussian HJM and Markovian HJM. In this chapter, we will introduce a new non-parametric method for calibrating Gaussian HJM models that overcomes the problems with standard approaches that we will shortly review.

Our aim is to calibrate Gaussian HJM to caps and swaptions. We show that it is more satisfactory to first calibrate the Gaussian Random Field Term Structure model of Kennedy (1994) before finding an HJM approximation to the calibrated Kennedy model. Our estimation is non-parametric in the sense that we do not have to assume any functional forms for the volatility factors.

Kennedy (1994) introduces a very general model where interest rates are modelled as a Gaussian Random Field. The Gaussian assumption allows for many closed form solutions to option pricing problems that are invaluable when we calibrate a model. Kennedy (1994) provides a cap pricing formula and we derive an approximate European Swaption pricing formula that can also be used for any of the existing Gaussian interest rate term structure models in the literature. The formulae allow us to calibrate Kennedy (1994) to a wide range of quoted caps and swaptions prices rapidly.

Kennedy (1994), however, does not allow for easy pricing of more sophisticated derivatives such as the path dependent variety. We can tackle such

pricing and hedging problems more easily in the HJM framework and it is for this reason that we use Kennedy as a device for calibrating HJM. With this in mind we focus on a particular specification of Kennedy (1994) that can be approximated easily by a multifactor Gaussian HJM model with time stationary volatility structures. This, we believe, provides for a better calibration of HJM models than by a direct calibration. We examine the fitted Kennedy model and find that quoted caps and swaptions prices produce two or possibly three significant factors. The approximating HJM models have time stationary volatility structures that our initial investigation suggests are also stable between re-calibration across calendar time.

We begin by giving a brief review of standard approaches and explaining how they are unsatisfactory in Section 7.1.1. We provide a review of Kennedy (1994) in Section 7.2 and examine the Kennedy covariance function we will fit in this paper in Section 7.3. Section 7.4 provides the key motivation factor for our choice of Kennedy covariance function; it shows how our chosen Kennedy model can be approximated by a stationary HJM model. Section 7.5 shows how contingent claims are priced in the Kennedy model and produces formulae that are used to calibrate the Kennedy model to caps and swaptions. We review the data we have employed in Section 7.6 and discuss the results and implications of the fitted model in Section 7.7. Section 7.8 summarises and makes a few concluding remarks.

7.1.1 Review of Conventional Approaches

7.1.1.1 Principal Components Analysis Calibration

Our review of interest rate pricing models in Chapter 2 noted that in the HJM approach, contingent claims can be priced once the forward rate volatility factors have been specified. Furthermore, the volatility factors are the same in the

risk-neutral measure as in the objective measure. This property suggests we can price contingent claims by extracting the volatility factors from historical time-series data. Some of the earliest calibration of HJM models extracted the volatility factors from historical data by conducting a Principal Components Analysis, (PCA), on a historically estimated covariance matrix of zero coupon yield changes or discount factor changes across a small time interval. We will describe briefly a PCA calibration of Gaussian HJM models. The PCA estimation of other forms of HJM models is similar. For convenience, we will assume the volatility factors are time-stationary.

The PCA approach proceeds as follows. Let $\Delta Y(t, t+\tau_i)$ denote the time t change of the τ_i maturity zero coupon yield across a small time interval Δt.

(1) Estimate

$$Cov[\Delta Y(t,t+\tau_i),\Delta Y(t,t+\tau_j)] = \widetilde{V}_{ij}(t)\Delta t \equiv \widetilde{V}_{ij}\Delta t \text{ for } i, j = 1 \text{ to } n$$

from historical time series data of zero coupon yield changes. The matrix $\widetilde{V}$ provides an estimate of V, the covariance matrix of instantaneous zero coupon yield changes for n distinct maturities.

(2) Calibrate a Gaussian HJM model to the time series data by defining the yield volatility factors implicitly by

$$dY(t,t+\tau_i) = \alpha_Y(t,t+\tau_i)dt + \sum_{k=1}^{n}\sqrt{\lambda_k}\,x_{ik}dW_k \qquad \textbf{(7.1)}$$

where $[x_{ik}]$ is a $(n \times n)$ matrix of the eigenvectors and $[\lambda_{ii}]$ is a $(n \times n)$ diagonal matrix of the corresponding eigenvalues of the matrix $\widetilde{V}$, that is, $\widetilde{V} = x\Lambda x^T$. $x_{ik}\sqrt{\lambda_k}$ is the loading of the kth volatility factor on the ith maturity zero coupon yield.

The HJM model given by equation 7.1 reproduces the sampled covariance matrix exactly for if we collect the chosen zero coupon yields into a column vector

$$d\underline{Y}(t,t+\tau) = \underline{\alpha}_Y(t,t+\tau)dt + x\sqrt{\Lambda}d\underline{W}$$

then

$$d\underline{Y}(t,t+\tau)d\underline{Y}(t,t+\tau)' = (x\Lambda x')dt = \tilde{V}dt\ .$$

Note that the procedure does not need to make an assumption for the number of factors that are driving the interest rate dynamics. A maximum of n is needed to reproduce the covariance matrix completely. It is easy to show that $\lambda_j / \Sigma_i\lambda_i$ is the proportion of the total variance accountable to jth factor.

Typical results point to the existence of three main factors accounting for about 90%, 5% and 1-2% of the total variance with each representing a parallel shift, change of slope and change in curvature, respectively, of the zero coupon yield curve. The results suggest two factor models should be able to capture interest rate dynamics very well and be able to price many interest rate derivatives accurately.

The PCA calibration we have described is very appealing because it is very intuitive. In practice, however, the resulting model fails to price consistently with observed market prices because of the misspecification problems we have already discussed in Chapter 2. For consistent pricing, the volatility factors have to be implied from options prices.

7.1.1.2 Implied Volatility Factors

We can improve the above procedure by using an implied fit, that is, we use market prices of options to extract the volatility factors. The implied approach will generally use an optimisation routine to find the model parameters that provide a good fit to market prices. Typically, the optimisation routine would may have to calculate the options prices for many different sets of trial parameters. The implied approach will therefore only be feasible if it possible to calculate options prices

quickly. This forces the practitioner to either restrict his choice of model and/or his choice of calibration options.

The Gaussian assumption we are focusing on in this chapter provides some analytical tractability. We can derive closed form pricing or approximation formulae for a modest number of options to enable a rapid calibration.

Sections 7.5.1 and 7.5.3 show we can obtain closed form solutions and approximations to the price of caps and swaptions respectively that depend only on the initial term structure and the covariance function $c(s, u, v)$ defined by

$$c(s_1 \wedge s_2, t_1, t_2) = Cov[F(s_1,t), F(s_2,t_2)] = \int_0^{s_1 \wedge s_2} \underline{\sigma}(u,t_1) \cdot \underline{\sigma}(u,t_2) du$$

where $F(u, t)$ denote the t maturity instantaneous forward rate at time u, $\underline{\sigma}(u, t)$ are the forward rate volatility factors and the $\cdot$ denotes the inner product. The caps and swaptions pricing formulae enable a simple implied calibration. Empirical data tells us what shapes the first three volatility factors take. So if we assume there are just three factors and make functional assumptions for their shapes, then we can extract the volatility factors from caps and swaptions prices.

Although this calibration procedure will enable the model to price the calibration options consistent with market prices, there are some problems with this approach. The functional forms assumed for the volatility factors have to be sufficiently flexible to produce similar shapes to the factors that can be extracted from historical data. For tractability, this suggests we use a high ordered polynomial, probably fourth. But in practice, the fitted polynomial volatility factors are unstable in the sense that the volatility factors can vary greatly when model is refitted to new market prices across time. This is clearly undesirable. Alternatively, we may approximate the volatility factors with splines. However, this makes the closed form cap pricing formula and swaption approximation

exceedingly difficult to evaluate analytically and a numerical integration too slow for calibration.

The problems we have discussed arise essentially from the indirect dependence of options prices on the volatility factors. We shall see in Sections 7.5.1 and 7.5.3 that the key unobserved variables for pricing options correspond to the volumes beneath the surface $c(s, u, v)$ for different values of s. These volumes are proportional to zero coupon yield covariances. Therefore, it should be more direct and simpler to fit zero coupon yield covariances rather than to fit zero coupon yield volatility factors that give the zero coupon yield covariances indirectly and only after a complicated integration. Fitting the zero coupon yield covariances allows the calibration to be separated into two distinct procedures. We first fit the zero coupon yield covariances and then find the zero coupon yield volatility factors that would reproduce the fitted zero coupon yield covariances. Note that it is unnecessary to assume, a priori, the number of factors driving the interest rate dynamics nor the shape of the volatility factors. The number of factors and their shapes will be determined when the zero coupon yield volatility factors are extracted from the zero coupon yield covariances. The remainder of this chapter shows how we can fit the zero coupon yield covariances and extract the zero coupon yield volatility factors.

7.2 REVIEW OF THE KENNEDY GAUSSIAN RANDOM FIELD TERM STRUCTURE MODEL

Unlike the majority of other interest rate modelling and derivative pricing models where interest rate processes are specified as stochastic differential equations for which one would have to solve to find the relevant risk-adjusted distributions, Kennedy (1994) instead specifies the risk-adjusted distributions such that no arbitrage opportunities exist. Kennedy (1994) requires as inputs the covariance of all future instantaneous forward rates and the initial instantaneous

forward rate curve. The joint distribution of all future forward rates in the risk-adjusted measure is then given by standard no-arbitrage arguments. Derivative prices are given simply by finding the appropriate expectations of the discounted payoffs under the risk-neutral measure. Kennedy's approach is complementary to the more traditional approach of specifying an interest rate process. Indeed Kennedy (1994) model encompasses all other diffusion models in which interest rates are normally distributed. Thus, for example, it is possible to find specifications for the Kennedy (1994) that corresponds to the models of Vasicek (1977), the Extended Vasicek model of Hull and White (1990), the multifactor model of Langetieg (1980) and any Gaussian HJM. Not only does the Kennedy model encompass all Gaussian interest rate models in the current literature, but it also allows for infinite factor models.

We adopt the notation used in Kennedy (1994) and use Kennedy to refer to Kennedy (1994) models. The model assumes

$$F(s, t) = \mu(s, t) + X(s, t), \tag{7.2}$$

$0 \le s \le t$, where $X(s, t)$ is a mean-zero continuous Gaussian random field. Kennedy (1995) shows the Gaussian and no arbitrage assumptions imply that $\mathrm{Cov}(F(s_1,t_1), F(s_2,t_2))$ must take the form $c(s_1 \wedge s_2, t_1, t_2)$, where $s_1 \wedge s_2$, denotes $\min(s_1, s_2)$, and that the random field has independent increments in the s-direction. That is, for any $0 \le s \le s' \le t$, the random variable $X(s', t) - X(s, t)$ is independent of the σ-field $\Im_s = \sigma\{X(u, v): u \le s, u \le v\}$. Furthermore, the absence of arbitrage implies that under the risk-neutral measure, the means of the future instantaneous forward rates must be related by

$$\mu(s,t) = \mu(0,t) + \int_0^t c(s \wedge v, v, t)dv \text{ for all } 0 \le s \le t. \tag{7.3}$$

Readers are referred to Kennedy (1994, 1995) for more details. Note that since zero coupon yields are averages of instantaneous forward rates, we can readily reformulate Kennedy (1994, 1995) in terms of zero coupon yields. Indeed, all the formulae and our calibration procedure in this paper can be adapted for the case when zero coupon yields are modelled as a Gaussian Random Field. To illustrate Kennedy (1994), we consider the following two examples.

7.2.1 Example 1: Hull-White Model

The risk-neutral process for the short rate is assumed to be

$$dr = \left[\phi(t) - a(t)r\right]dt + \sigma(t)dz$$

where $\phi(t)$, $a(t)$ and $\sigma(t)$ are deterministic functions of time usually chosen such that the model is made consistent with an observed zero coupon yield curve, a zero coupon yield volatility term structure and perhaps prior believes on short rate volatilities. Hull and White (1990) show that under the risk-neutral process, the time t price of a pure discount bond with maturity T is given by

$$P(t,T) = A(t,T)\exp[-B(t,T)r(t)]$$

where

$$B(t,T) = \frac{B(0,T) - B(0,t)}{\partial B(0,t)/\partial t}$$

$$B(0,t) = \frac{R(0,t)\sigma_R(0,t)t}{\sigma(0)}$$

$\sigma_R(0,t)$ is the proportional volatility of a t-maturity zero coupon yield at time 0, $R(0, t)$ is the t-maturity zero coupon yield at time 0, $\sigma(0)$ is the short rate proportional volatility at time 0 and $A(t, T)$ is a deterministic function of the initial term structure given by

$$\ln A(t,T) = \ln\frac{A(0,T)}{A(0,t)} - B(t,T)\frac{\partial}{\partial t}\ln A(0,t) - \frac{1}{2}\left[B(t,T)\frac{\partial B(0,t)}{\partial t}\right]^2 \int_0^t \left[\frac{\sigma(\tau)}{\partial B(0,\tau)/\partial \tau}\right]^2 d\tau .$$

It is shown in Appendix 1 that the covariance between the two instantaneous forward rates $F(s_1,t_1)$ and $F(s_2,t_2)$ is given by

$$Cov[F(s_1,t_1),F(s_2,t_2)] = \frac{\partial B(0,t_1)}{\partial t_1}\frac{\partial B(0,t_2)}{\partial t_2}\int_0^{s_1\wedge s_2}\left[\frac{\sigma(u)}{\partial B(0,u)/\partial u}\right]^2 du .$$

Thus the Hull and White (1990) model can be put into the Kennedy framework by specifying that

$$c(s_1\wedge s_2,\ t_1,\ t_2) = f(s_1\wedge s_2)g(t_1,\ t_2) \tag{7.4}$$

with the functions $f(s)$ and $g(t_1,\ t_2)$ defined by equations 7.5 and 7.6 respectively.

$$f(s) = \int_0^s\left[\frac{\sigma(u)}{\partial B(0,u)/\partial u}\right]^2 du \tag{7.5}$$

$$g(t_1,t_2) = \frac{\partial B(0,t_1)}{\partial t_1}\frac{\partial B(0,t_2)}{\partial t_2} . \tag{7.6}$$

7.2.2 Example 2: Gaussian Heath, Jarrow and Morton

The risk-neutral instantaneous forward rate process is given by

$$dF(s,u) = \alpha(s,u)dt + \underline{\sigma}(s,u)\cdot d\underline{z}$$

where $s \le u$, $\underline{\sigma}(s,\ u)$ is a n element column vector of deterministic volatility factors and $d\underline{z}$ is a n element column vector of independent Brownian increments. HJM (1992) shows that $\alpha(s,\ u)$ is constrained by no-arbitrage arguments to be given by

$$\alpha(s,u) = \underline{\sigma}(s,u)\cdot\int_s^u \underline{\sigma}(s,v)dv .$$

Assuming that the volatility factors satisfy the appropriate conditions of HJM (1992), then the instantaneous forward rate $F(s,\ t)$ is given by

$$F(s,t) = F(0,t) + \int_0^s \alpha(u,t)du + \int_0^s \underline{\sigma}(u,t)\cdot d\underline{z}(u) .$$

It follows that the equivalent specification for Kennedy follows immediately since

$$Cov[F(s_1,t_1),F(s_2,t_2)] = c(s_1 \wedge s_2,t_1,t_2) = \int_0^{s_1 \wedge s_2} \underline{\sigma}(u,t_1) \cdot \underline{\sigma}(u,t_2)du. \tag{7.7}$$

We have seen how all Gaussian HJM models can be put into the Kennedy framework. Equation 7.7 also suggests we can approximate Kennedy with a finite factor Gaussian HJM. We show how this can be achieved easily for a particular class of Kennedy covariance functions in Section 7.4. Before that we specify our non-parametric covariance function and provide sufficient conditions for the continuity and smoothness of the term structure of forward rates.

7.3 COVARIANCE FUNCTION, CONTINUITY AND SMOOTHNESS

There are various ways in which one can utilise market data to calibrate interest rate derivative pricing models. In the context of Kennedy, we need to ascertain what the covariance structure of future instantaneous rates are. Given the current interest rate term structure, the means are then determined by standard no-arbitrage arguments. We need to estimate the covariance function $c(s, u, v)$ and ensure that the fitted model will be well-behaved. This section examines the conditions that will ensure that the term structure of the forward rates will be continuous and smooth.

As we explained in the introduction, we want to find a Kennedy model that can be readily approximated by a finite factor Gaussian HJM model. We will assume that the volatility structures are time stationary and discuss how to allow for time-dependence later.

We first derive the functional form that $c(s, u, v)$ must have to provide time-stationary covariance rates. Consider

$$Cov[dF(s,t_1),dF(s,t_2)|\Im_s] = \lim_{t \to s^+} \frac{Cov\left[\{F(t,t_1)-F(s,t_1)\},\{F(t,t_2)-F(s,t_2)\}|\Im_s\right]}{t-s} ds$$

$$= \lim_{t \to s^+} \frac{Cov\left[F(t,t_1), F(t,t_2) | \Im_s\right]}{t-s} ds$$

$$= \lim_{t \to s^+} \frac{c(t,t_1,t_2) - c(s,t_1,t_2)}{t-s} ds$$

$$= \frac{\partial}{\partial s^+} c(s,t_1,t_2) ds . \qquad \textbf{(7.8)}$$

It follows from equation 7.8 that for time stationary covariance rates we must have

$$\frac{\partial}{\partial s^+} c(s,t_1,t_2) = g(t_1 - s, t_2 - s)$$

for some non-negative definite function $g(u, v)$. Integrating gives

$$c(s,t_1,t_2) = \int_0^s g(t_1 - u, t_2 - u) du + h(t_1,t_2) .$$

We therefore assume that the Kennedy covariance function of this paper satisfies:

Assumption 1:

$c(s,t_1,t_2) = \int_0^s g(t_1 - u, t_2 - u) du$ *for all* t_2, $t_1 \geq s \geq 0$ *with* $g(u, v)$ *satisfying the conditions of Proposition 1.*

We assume $h(t_1, t_2) = 0$ for all t_1 and t_2 so that $c(0, t_1, t_2) = 0$ for all t_1 and t_2. This corresponds to assuming that $\Im_0$ is the trivial σ-field. Kennedy (1995) examines interesting cases where $\Im_0$ is non-trivial. We call the covariance function of Assumption 1 the stationary Kennedy covariance function. Proposition 1 provides a sufficient condition for the continuity of the forward rate term structure.

Proposition 1: *For* $Cov(F(s_1,t_1),F(s_2,t_2)) = c(s_1 \wedge s_2, t_1, t_2) = \int_0^{s_1 \wedge s_2} g(t_1 - u, t_2 - u)du$, $0 \le s_1 \le t_1 < \infty$, $0 \le s_2 \le t_2 < \infty$, *the Gaussian Random Field has continuous sample functions, with probability one, when* $g(u, v)$ *is continuous and bounded.*

Proposition 1 is proved in Appendix 2. The forward rate surface will be continuous when $g(u, v)$ of the stationary Kennedy covariance function is continuous and bounded. The conditions of Proposition 1 are very natural so they impose no significant constraints on the type of covariance functions we may want to use. Proposition 2 provides a sufficient condition for the smoothness of the forward rate term structure.

Proposition 2: *For* $Cov(F(s_1,t_1),F(s_2,t_2)) = c(s_1 \wedge s_2, t_1, t_2) = \int_0^{s_1 \wedge s_2} g(t_1 - u, t_2 - u)du$, *for finite* s_1, s_2, t_1 *and* t_2, *the term structure of instantaneous forward rates is smooth, with probability one, when* $\frac{\partial^2 g(\tau_1,\tau_2)}{\partial\tau_1\partial\tau_2}$ *is continuous and bounded.*

Proof: Consider

$$Cov\left[\frac{dF(s_1,t_1)}{dt_1}, \frac{dF(s_2,t_2)}{dt_2}\right]$$

$$= \lim_{\Delta t_1 \to 0} \lim_{\Delta t_2 \to 0} Cov\left[\frac{F(s_1,t_1+\Delta t_1) - F(s_1,t_1)}{\Delta t_1}, \frac{F(s_2,t_2+\Delta t_2) - F(s_2,t_2)}{\Delta t_2}\right]$$

$$= \lim_{\Delta t_1 \to 0} \frac{1}{\Delta t_1} \lim_{\Delta t_2 \to 0}\left[\frac{c(s_1 \wedge s_2, t_1+\Delta t_1, t_2+\Delta t_2) - c(s_1 \wedge s_2, t_1+\Delta t_1, t_2)}{\Delta t_2} - \frac{c(s_1 \wedge s_2, t_1, t_2+\Delta t_2) - c(s_1 \wedge s_2, t_1, t_2)}{\Delta t_2}\right]$$

$$= \lim_{\Delta t_1 \to 0} \frac{1}{\Delta t_1}\left[\frac{\partial c(s_1 \wedge s_2, t_1+\Delta t_1, t_2)}{\partial t_2} - \frac{\partial c(s_1 \wedge s_2, t_1, t_2)}{\partial t_2}\right]$$

$$= \frac{\partial^2 c(s_1 \wedge s_2, t_1, t_2)}{\partial t_1 \partial t_2} = \int_0^{s_1 \wedge s_2} \frac{\partial^2 g(t_1 - u, t_2 - u)}{\partial t_1 \partial t_2} du .$$

The covariance function has a similar functional form to that of Proposition 1. Therefore the proof of Proposition 1 can be used here also to provide the result that $\frac{dF(s,t_1)}{dt_1}$ is continuous when $\frac{\partial^3 c(s,t_1,t_2)}{\partial s \partial t_1 \partial t_2} = \frac{\partial^2 g(t_1 - s, t_2 - s)}{\partial t_1 \partial t_2}$ is continuous and bounded. **QED**.

Propositions 1 and 2 provide constraints on the function $g(u, v)$ that we are going to fit to caps and swaptions prices. There is a compromise between the ease with which $g(u, v)$ can be fitted and an ideal. With the aim of simplifying the estimation of the covariance function we make the following assumption:

Assumption 2: *$g(u, v)$ can be approximated by a symmetric and piecewise-triangular surface: For node times t_i, $i = 1, \ldots, n$, $t_i < t_j$ for $i < j$, the four corners (t_i, t_j), (t_{i+1}, t_j), (t_i, t_{j+1}) and (t_{i+1}, t_{j+1}) of the approximating surface define two piecewise triangular sections that are joined along the line running from (t_i, t_j) to (t_{i+1}, t_{j+1}).*

Assumption 2 allows us to calibrate the stationary Kennedy model much more quickly than otherwise. It is non-parametric and sufficiently flexible to approximate any continuous surface well. However, it does not meet the criteria of Proposition 2 for a smooth term structure of forward rates. However, this is more of a theoretical rather than a practical problem for we can assume that the edges where the triangular planes of $g(u, v)$ meet are rounded off to provide continuous and bounded second derivatives. We will see in Section 7.5 that the rounding off of the edges have little effect on caps and swaptions prices since those prices

depend on the volume beneath $g(u, v)$. The HJM approximation prescribed in Section 7.4 will give the rounded edges.

We calibrate the Kennedy covariance function of Assumption 1 to market caps and swaptions prices in Section 7.7. Our optimisation imposes the condition that the matrix $G = [g(t_i, t_j)]_{i,j=1..n}$ be positive definite. We do not know whether the interpolation rule of Assumption 2 and the positive definiteness of G are sufficient conditions for the positive definiteness of the fitted function $g(u, v)$. The function $g(u, v)$ generated by the approximating HJM model of Section 7.4 will however be positive definite.

It is arguable that our time stationary covariance function is unsatisfactory because empirical data suggests that volatilities do vary through time. We can address this issue if the covariance function is assumed to be

$$c(s,t_1,t_2) = \int_0^s f(u)g(t_1 - u, t_2 - u)du \tag{7.9}$$

where $f(s)$ is a non-negative function. Then it follows from equation 7.8 that

$$Cov\left[dF(s,t_1), dF(s,t_2) \middle| \Im_s\right] = f(s)g(t_1 - s, t_2 - s)\,. \tag{7.10}$$

The covariance function of equation 7.9 allow for time-dependent volatilities. The covariance surface, equation 7.10, as a function of the times to maturities, maintain its shape through time but its level is permitted to shift up and down to reflect changing volatility levels.

It is useful to contrast equation 7.10 with the corresponding expression for the Gaussian Hull and White (1990). Using equations 7.4 to 7.6 we can show that the covariance rates are given by

$$\text{HW(1990):}\quad Cov\left[dF(s,t_1), dF(s,t_2) \middle| \Im_s\right] = h(s)\frac{\partial B(0,t_1)}{\partial t_1}\frac{\partial B(0,t_2)}{\partial t_2} \tag{7.11}$$

where $h(s)$ is the right derivative of equation 7.5 with respect to s. Equation 7.11 shows that, in the Gaussian Hull and White (1990) model, the shape of the

covariance surface, as a function of the times to maturities, evolves in a very unsatisfactory way through time. In effect, we ride up on the initial covariance surface as time passes. This is the problem with the Hull and White (1990) model we discussed in Chapter 4.

The covariance function of equation 7.9 provides realistic behaviour. It would allow a better fit to the market prices than the time stationary covariance function of Assumption 1. Our results in section 7.7 show that the time-stationary covariance function already allows the Kennedy model to fit caps and swaptions prices accurately so we do not fit the covariance function of equation 7.9.

7.4 IMPLIED COVARIANCE OF ZERO COUPON YIELD CHANGES AND HJM CALIBRATION

After fitting the Kennedy model to market caps and swaptions prices, it will be important to examine the implied covariances of zero coupon yield changes to see whether they are plausible. This section derives a formula for the implied covariances and also explains how to extract the volatility structures for the HJM approximation to the fitted Kennedy model.

The t-s maturity zero coupon yield at time s is given by

$$Y(s,t) = \frac{1}{t-s}\int_s^t F(s,u)du.$$

For any t_1, $t_2 \geq s$ we have

$$Cov\left[dY(s,t_1), dY(s,t_2)\middle|\Im_s\right] = \frac{1}{(t_1-s)(t_2-s)} Cov\left[\int_s^{t_1} dF(s,u)du, \int_s^{t_2} dF(s,v)dv\middle|\Im_s\right]$$

$$= \frac{1}{(t_1-s)(t_2-s)}\int_s^{t_1}\int_s^{t_2} Cov[dF(s,u), dF(s,v)|\Im_s]dvdu$$

$$= \frac{1}{(t_1-s)(t_2-s)}\int_s^{t_1}\int_s^{t_2} \frac{\partial}{\partial s^+} c(s,u,v)dvdu$$

$$= \frac{1}{(t_1 - s)(t_2 - s)} \int_s^{t_1} \int_s^{t_2} g(u-s, v-s)\,dv\,du$$

$$= \frac{1}{\tau_1 \tau_2} \int_0^{\tau_1} \int_0^{\tau_2} g(u,v)\,dv\,du \qquad \textbf{(7.12)}$$

where $\tau_1 = t_1 - s$ and $\tau_2 = t_2 - s$ are the maturities of the two zero coupon yields considered. The volatility term structure is time stationary. Our choice of the covariance function is motivated by this result. It also allows for a very simple HJM approximation to the fitted Kennedy model. Suppose we want to construct the HJM volatility factors at n distinct maturities, τ_i, $i = 1..n$, $\tau_i > \tau_j$ if $i > j$. Then the HJM approximation is characterised by

$$\sigma_i(\tau_j) = \sqrt{\Lambda_{ii}}\, M_{ji}$$

where $\sigma_i(\tau_j)$ is the ith stationary volatility factor for a τ_j maturity zero coupon yield; Λ and M are respectively the eigenvalues and eigenvectors of the covariance matrix C where

$$C_{ij} = Cov\left[dY(0,\tau_i), dY(0,\tau_j) \middle| \Im_0\right] = \frac{1}{\tau_i \tau_j} \int_0^{\tau_i} \int_0^{\tau_j} g(u,v)\,dv\,du\,. \qquad \textbf{(7.13)}$$

The volatility factors can be completed by interpolating, for example using cubic-splines, between the chosen maturities and the accuracy of the approximation improved by increasing n. Our approach is related to the use of PCA in some popular implementations of Gaussian HJM models. Our approach conducts a PCA on the implied covariance matrix of zero coupon yield changes whereas in more common approaches, a PCA is conducted on a covariance matrix of historical zero coupon yield changes. The latter implementations have not been successful at matching market cap and swaption prices as discussed earlier. Our calibration procedure provides volatility structures that are implied from market caps and swaptions prices. Typically, we can choose a large number for n to ensure the

HJM approximation is good. No numerical tractability is lost since the HJM approximation will typically have at most three significant factors. We believe our procedure is preferable to working with the HJM model directly. We do not have to pre-specify the functional form for the volatility factors and do not have to pre-specify the number of volatility factors required.

Before we consider derivatives pricing, we note that if we wanted time-dependence and used the covariance function given by equation 7.9 instead of the time-stationary covariance function, then it follows from equation 7.12 that

$$Cov\left[dY(s,t_1),dY(s,t_2)|\Im_s\right] = \frac{f(s)}{\tau_1\tau_2}\int_0^{\tau_1}\int_0^{\tau_2} g(u,v)dvdu$$

where τ_1 and τ_2 are defined as before. The zero coupon yield volatility factors maintain their shapes through time but the volatility levels are permitted to shift up and down to reflect changing volatility levels. The volatility term structures are still given by the eigenvectors of the matrix C defined by equation 7.13 and the approximating HJM model is characterised by

$$\sigma_i(s,\tau_j) = \sqrt{f(s)\Lambda_{ii}}\,M_{ji}$$

where $\sigma_i(s, \tau_j)$ is the ith volatility factor for a τ_j zero coupon yield at time s.

7.5 THE PRICING OF CONTINGENT CLAIMS

As shown by Geman, El Karoui and Rochet (1995) and Jamshidian (1989), it is now well known that for the purpose of deriving derivative pricing formulae, a change of numéraire or probability measure can simplify the process greatly. For the pricing of caps and swaptions, it is convenient to use a suitably chosen pure discount bond, (PDB), as the numéraire. The standard results of Harrison and Kreps (1979) and Harrison and Pliska (1981) then allow us to use the equivalent martingale measure (EMM) that renders prices measured in our chosen numéraire a martingale to price contingent claims consistently.

We use $P(s, t)$ to denote the time s price of a PDB which matures at time t. When the chosen numéraire is a t maturity PDB, we call the EMM the t-measure and use $E^t[]$ to indicate expectations taken with respect to the t-measure. Note that changing the numéraire does not change the forward rate covariances. Therefore the Kennedy covariance function is invariant in the following propositions.

Proposition 3: *The following two statements are equivalent:*

a) *For the chosen numéraire,* $P(s,t_1)$, *the process* $\{P(s,t_2)/P(s,t_1),\ \Im_s,\ 0 \le s \le t_1 \le t_2\}$ *is a martingale;*

b) $\mu(s_1,t_2) = \mu(s,t_2) + \int_{t_1}^{t_2} \{c(s_1,t_2,v) - c(s,t_2,v)\}dv$ *for* $0 \le s \le s_1 \le t_1 \le t_2$.

Proof: We have from the definition of forward rates that

$$\frac{P(s_1,t_2)}{P(s_1,t_1)} = \frac{P(s,t_2)}{P(s,t_1)} \exp\left\{-\int_{t_1}^{t_2} \{F(s_1,u) - F(s,u)\}du\right\} \text{ where } s \le s_1$$

so that

$$E^{t_1}\left[\frac{P(s_1,t_2)}{P(s_1,t_1)}\middle|\Im_s\right] = \frac{P(s,t_2)}{P(s,t_1)} E^{t_1}\left[\exp\left(-\int_{t_1}^{t_2}\{F(s_1,u) - F(s,u)\}du\right)\middle|\Im_s\right]$$

$$= \frac{P(s,t_2)}{P(s,t_1)} E^{t_1}\left[\exp\left(-\int_{t_1}^{t_2}\{F(s_1,u) - F(s,u)\}du\right)\right]$$

where the second equality follows since the random variable $\{F(s_1,u) - F(s,u),\ s \le s_1 \le u\}$ is independent of $\Im_s$. Now since

$$E^{t_1}\left[-\int_{t_1}^{t_2}\{F(s,u) - F(s_1,u)\}du\right] = \int_{t_1}^{t_2}\{\mu(s_1,u) - \mu(s,u)\}du$$

and

$$Var\left[\int_{t_1}^{t_2}\{F(s_1,u)-F(s,u)\}du\right]=\int_{t_1}^{t_2}\int_{t_1}^{t_2}\{c(s_1,u,v)-c(s,u,v)\}dvdu$$

it follows that the process $\{P(s,t_2)/P(s,t_1),\ \Im_s,\ 0\le s\le t_1\le t_2\}$ is a martingale when

$$\int_{t_1}^{t_2}\{\mu(s_1,u)-\mu(s,u)\}du=\int_{t_1}^{t_2}\int_{t_1}^{u}\{c(s_1,u,v)-c(s,u,v)\}dvdu$$

since $c(.,\ u,\ v)$ is symmetric in u and v. Differentiating with respect to t_2 gives (*b*).

Now consider $\dfrac{P(s_1,t_2)}{P(s_1,t_1)}=\exp\left[-\int_{t_1}^{t_2}F(s_1,u)du\right]$. Using (b) we have that

$$E^{t_1}\left[\int_{t_1}^{t_2}F(s_1,u)du\middle|\Im_s\right]=\int_{t_1}^{t_2}F(s,u)du+\int_{t_1}^{t_2}\int_{t_1}^{u}\{c(s_1,u,v)-c(s,u,v)\}dvdu$$

and

$$Var\left[\int_{t_1}^{t_2}F(s_1,u)du\middle|\Im_s\right]=\int_{t_1}^{t_2}\int_{t_1}^{t_2}\{c(s_1,u,v)-c(s,u,v)\}dvdu$$

so that

$$E^{t_1}\left[\frac{P(s_1,t_2)}{P(s_1,t_1)}\middle|\Im_s\right]=\exp\left\{-\int_{t_1}^{t_2}F(s,u)du\right\}=\frac{P(s,t_2)}{P(s,t_1)}.$$

QED.

Proposition 4: *The following two statements are equivalent:*

a) *For the chosen numéraire, $P(s,t_2)$, the process $\{P(s,t_1)/P(s,t_2),\ \Im_s,\ 0\le s\le t_1\le t_2\}$ is a martingale;*

b) $\mu(s_1,t_1)=\mu(s,t_1)-\int_{t_1}^{t_2}\{c(s_1,u,t_1)-c(s,u,t_1)\}du$ *for* $0\le s\le s_1\le t_1\le t_2$.

Proof: The proof follows from the same type of calculations as that of Proposition 3. **QED**.

7.5.1 Pricing of Caps

Kennedy (1994) derives the time $s \le t$ price of a caplet that has payoff at time $t+\delta$,

$$[\exp(\delta Y(t,t+\delta)) - \exp(\delta k)]^+,$$

where

$$Y(t,t+\delta) = \frac{1}{\delta}\int_t^{t+\delta} F(t,u)du$$

is the δ maturity zero coupon yield at time t and k is the cap rate by taking expectations with respect to the risk-neutral measure. We derive the same formula to illustrate that a well chosen numéraire can simplify the calculations greatly. We chose as numéraire the PDB with maturity $t+\delta$. Thus the value of the caplet at time $s \le t$ is given by the expectation

$$cpt(s) = P(s,t+\delta)E^{t+\delta}\left[\left(\exp\left[\int_t^{t+\delta} F(t,v)dv\right] - \exp(k\delta)\right)^+ \middle| \Im_s\right].$$

Proposition 4 gives

$$\mu(t,u) = \mu(s,u) - \int_u^{t+\delta}\{c(t,v,u) - c(s,v,u)\}dv$$

so that

$$E^{t+\delta}\left[\int_t^{t+\delta} F(t,u)du \middle| \Im_s\right] = \int_t^{t+\delta} F(s,u)du - \int_t^{t+\delta}\int_u^{t+\delta}\{c(t,v,u) - c(s,v,u)\}dvdu$$

$$= \int_t^{t+\delta} F(s,u)du - \int_t^{t+\delta}\int_t^{u}\{c(t,v,u) - c(s,v,u)\}dvdu$$

where the second equality follows since $c(\cdot, v, u)$ is symmetric in u and v and

$$Var\left[\int_t^{t+\delta} F(t,u)du \middle| \Im_s\right] = \int_t^{t+\delta}\int_t^{t+\delta}\{c(t,v,u) - c(s,v,u)\}dvdu.$$

Taking expectations now gives

$$cpt(s) = P(s,t)N(d_1) - P(s,t+\delta)e^{\delta k}N(d_2)$$

where

$$d_1 = \frac{\ln\left[\frac{P(s,t)}{P(s,t+\delta)}\right] - k\delta}{\sigma} + \frac{\sigma}{2},\ d_2 = d_1 - \sigma$$

and

$$\sigma^2 = \int_t^{t+\delta}\int_t^{t+\delta}\{c(t,v,u) - c(s,v,u)\}dvdu.$$

Note that the market convention in London is to quote rates on a quarterly basis with payoff at time $t + \frac{1}{4}$ of $\frac{1}{4}\left[f_t^{t,t+\frac{1}{4}} - k\right]^+$ where $f_t^{t,t+\frac{1}{4}}$ is the floating rate for borrowing or lending over the period $[t, t + \frac{1}{4}]$ and k the cap rate both expressed quarterly. However since

$$f_t^{t,t+\frac{1}{4}} = 4\left[\exp\left(\frac{F_t^{t,t+\frac{1}{4}}}{4}\right) - 1\right]$$

then denoting k' for the cap rate in continuous terms we have that the

$$\text{payoff} = \left[\exp\left(\frac{F_t^{t,t+\frac{1}{4}}}{4}\right) - \exp\left(\frac{k'}{4}\right)\right]^+$$

with $k' = 4\ln(1 + k/4)$. Thus we can use the caplet formula provided we convert the cap rate to continuous compounding.

7.5.2 Pricing of European Swaptions

In the London market, the settlement value for an European Payer Swaption, with maturity s, on an n year USD swap is given by

$$\left[w(s)-k\right]^{+}\sum_{i=1}^{n}\frac{1}{\left[1+w(s)\right]^{i}} \tag{7.14}$$

where $w(s)$ is the swap rate at time s for an n year USD swap and k is the strike rate. We do not have an exact closed form formula for the price of the swaption above but the next proposition allows us to estimate the price efficiently by numerically evaluating the expectation using Quasi Monte Carlo Methods as in Joy, Boyle and Tan (1995). Proposition 5 allows us to test the accuracy of the approximate European swaption pricing formula of Proposition 6. Let the time s_i denote $s + i$.

Proposition 5: *The time u price of an European payer swaption with maturity $s \geq u$ and strike k on a n year swap with the payoff*

$$\left[w(s)-k\right]^{+}\sum_{i=1}^{n}\frac{1}{\left[1+w(s)\right]^{i}}$$

is given by

$$Swpt(u) = P(u,s)E^{s}\left[\left[w(s)-k\right]^{+}\sum_{i=1}^{n}\frac{1}{\left(1+w(s)\right)^{i}}\middle|\Im_{u}\right] \tag{7.15}$$

where

$$w(s) = \frac{1-P(s,s_n)}{\sum_{j=1}^{n}P(s,s_j)}; \tag{7.16}$$

Defining X_i such that

$$P(s,s_i) = exp(X_i), \tag{7.17}$$

then under the s-measure, the random variable $\{X_i \mid \Im_u\}$ is normally distributed with mean μ_i and variance σ_i^2 given by

$$\mu_i = \ln\left[\frac{P(u,s_i)}{P(u,s)}\right] - \frac{1}{2}\sigma_i^2 \tag{7.18}$$

and

$$\sigma_i^2 = \int_s^{s_i}\int_s^{s_i}\{c(s,a,b) - c(u,a,b)\}dbda\,. \tag{7.19}$$

Moreover

$$Corr(X_i, X_j \mid \Im_u) = \frac{\int_s^{s_i}\int_s^{s_j}\{c(s,a,b) - c(u,a,b)\}dbda}{\sqrt{\int_s^{s_i}\int_s^{s_i}\{c(s,a,b) - c(u,a,b)\}dbda\int_s^{s_j}\int_s^{s_j}\{c(s,a,b) - c(u,a,b)\}dvdu}}\,. \tag{7.20}$$

Proof: It is well known that the swap rate on an n year USD swap at time s would be given by

$$w(s) = \frac{1 - P(s,s_n)}{\sum_{j=1}^{n} P(s,s_j)}\,. \tag{7.21}$$

It follows that, under the s-measure induced by taking the time s maturity PDB as the numéraire, the value of the payer swaption at any time $u \le s$ is given by the expectation of equation 7.15. The distribution of $\{P(s, s_i) \mid \Im_u\}$ and hence $\{X_i \mid \Im_u\}$ follows from using Proposition 3 and simplifying to give equations 7.18 and 7.19. Equation 7.20 follows from using

$$Cov[X_i, X_j|\Im_u] = \int_s^{s_i}\int_s^{s_j}\{c(s,a,b) - c(u,a,b)\}dbda\,.$$

QED.

The proposition allows us to value European Payer Swaptions using

$$Swpt(0) = \sum_{j=1}^{N} Swpt_j(0)$$

where the jth sample value is given by

$$Swpt_j(0) = P(0,s)\left[w(s) - k\right]^+ \sum_{i=1}^{n} \frac{1}{\left[1 + w(s)\right]^i}$$

with $w(s)$ of equation 7.16 calculated using

$$\left\{P(s,s_i)\middle|\Im_0\right\} = \exp\left[\mu_i + \sigma_i \varpi_i(j)\right]$$

where μ_i and σ_i are given by equations 7.18 and 7.19 respectively, $\varpi_i(j)$ is the ith element of $\underline{\varpi}(j)$ which is the jth realisation of a random vector of standard normal random variates with $Corr(\varpi_i, \varpi_k) = Corr(X_i, X_k \mid \Im_0)$ as given by equation 7.20, and N is the total number of sample values. We generate the correlated normal variates, $\underline{\varpi}(j)$, using the transformation $\underline{\varpi}(j) = M\Lambda^{1/2}\underline{\varepsilon}(j)$ where $\underline{\varepsilon}(j)$ is a sample random vector of uncorrelated standard normal variates, and M and Λ are respectively matrices with the eigenvectors and eigenvalues of the required correlation matrix. $\Lambda^{1/2}$ is a diagonal matrix with the elements set equal to the square root of the respective elements in Λ. The standard uncorrelated normal variates are generated from an n-dimensional Faure sequence. Joy, Boyle and Tan (1995) shows that the estimate has an error of $O(1/N)$.

7.5.3 An Approximate European Swaption Pricing Formula

To derive our approximate formula we make two assumptions. The first assumes, at maturity s, the payer swaption has payoff given by

$$\left[w(s) - k\right]^+ \sum_{i=1}^{n} P(s,s_i) \qquad \textbf{(7.22)}$$

instead of the USD market convention of equation 7.14 where $w(s)$ is the swap rate for an n year USD swap and $s_i = s + i$. Equation 7.14 discounts future cashflows with the time s par yield of an n year coupon bond whereas equation 7.22 discounts the cashflows with the time s zero coupon yields. Note that in

some markets, swaption payoffs are defined by equation 7.22 so this approximation would not be necessary and the swaption approximation formula we will derive will be more accurate. With this simplifying assumption, substituting equation 7.21 into equation 7.22 and simplifying give the payoff as

$$\left[1-\sum_{i=1}^{n} k_i P(s,s_i)\right]^+$$

where $k_i = k$ for $i = 1..n\text{-}1$ and $k_n = 1+k$. Thus the payoff to the payer swaption is the same as that of a put option on a coupon bond with strike one. We next assume that the underlying coupon bond at the maturity of the swaption is lognormally distributed. The two assumptions allow Proposition 6:

Proposition 6: *An approximate formula for the time u price of an European Payer Swaption with maturity s ≥ u and strike rate k on an n year swap with the maturity payoff*

$$[w(s)-k]^+ \sum_{i=1}^{n} \frac{1}{[1+w(s)]^i}$$

is given by

$$P(u,s)\left\{N\left[-\frac{a}{b}\right]-e^{a+\frac{b^2}{2}}N\left[-\frac{a+b^2}{b}\right]\right\} \tag{7.23}$$

where

$$a=\ln\left[\frac{m}{\sqrt{1+v/m^2}}\right], \tag{7.24}$$

$$b^2=\ln\left[1+v/m^2\right], \tag{7.25}$$

$$m=\sum_{i=1}^{n}\frac{k_i P(u,s_i)}{P(u,s)}, \tag{7.26}$$

$$v = \underline{K}^T C \underline{K}, \tag{7.27}$$

$$\underline{K}^T = [k \;\; k \;\; \cdots \;\; k \;\; 1+k] \tag{7.28}$$

k_i is the ith element of the vector $\underline{K}$ and C is an n x n matrix with the element (i, j) given by

$$C_{ij} = \frac{P(u,s_i)P(u,s_j)}{P(u,s)^2}\left\{\exp\left(\int_s^{s_i}\int_s^{s_j}\{c(s,x,y)-c(u,x,y)\}dydx\right)-1\right\}. \tag{7.29}$$

Proof: Under the s-measure and with our simplifying assumptions, the time u price of the swaption is given by

$$payer(u) = P(u,s)E^s\left[\left(1-\sum_{i=1}^{n}k_i P(s,s_i)\right)^+\middle|\Im_u\right].$$

Under the s-measure, we have

$$E^s[P(s,s_i)|\Im_u] = \frac{P(u,s_i)}{P(u,s)}$$

$$Cov\left[P(s,s_i),P(s,s_j)\middle|\Im_u\right] = \frac{P(u,s_i)P(u,s_j)}{P(u,s)^2}\left\{\exp\left(\int_s^{s_i}\int_s^{s_j}\{c(s,a,b)-c(u,a,b)\}dbda\right)-1\right\}$$

so that the mean and variance of the coupon bond at time s is given by m and v, defined by equations 7.26 and 7.27 respectively. As the coupon bond price is assumed to be lognormal, we can write

$$\left\{\sum_{i=1}^{n}P(s,s_i)\middle|\Im_u\right\} = e^Z, Z \sim N(a,b^2).$$

Matching the means and variance of the distribution gives a and b defined by equations 7.24 and 7.25 respectively. Finally evaluating $P(u,s)E^s\left[(1-e^Z)^+|\Im_u\right]$ gives the approximate formula, equation 7.23. **QED**.

Corollary: *An approximate formula for the time u price of an European Receiver Swaption with maturity s ≥ u and strike k on a n year swap with the maturity payoff*

$$[k - w(s)]^+ \sum_{i=1}^{n} \frac{1}{[1 + w(s)]^i}$$

is given by

$$P(u,s)\left\{e^{a+\frac{b^2}{2}} N\left[\frac{a+b^2}{b}\right] - N\left[\frac{a}{b}\right]\right\}$$

where a and b are as defined in Proposition 6.

Proof: Follows from recognising that the receiver swaption is given by

$$P(u,s)E^s\left[(e^Z - 1)^+ \middle| \Im_u\right]$$

with Z as defined in Proposition 6, and that the expectation can be obtained by making the same changes one makes to an European Put formula to obtain an European Call on a lognormal underlying. Alternatively, evaluate the expectations directly. **QED**.

The approximate swaption pricing formula of Proposition 6 allows us to calibrate Kennedy far more quickly than would have been possible using a numerical integration. The approximation performs very well for close to the money swaptions. For typical covariance structures, the differences between approximation and the numerical integration are very small. To illustrate this, we use the implied covariance function of 31st May 1996 in Section 7.7, shown in Table 5. Table 1 compares the European Payer Swaption prices produced by a numerical integration of Proposition 5 with those by the approximation formula of Proposition 6. The numerical integration uses one million samples so the errors are of $O(10^{-6})$. The approximation formula of Proposition 6 performs well for near the money swaptions. This result is not unexpected as the distribution of the

coupon bond is the convolution of a number of correlated lognormal distributions. For small coupons, the present value of the principal dominates and so the distribution of the coupon bond would be approximately lognormal. We can conclude that since quoted swaptions are at-the-money, the approximation formula would be appropriate for the calibration in Section 7.7. Note that the approximation would perform even better for swaptions with maturity payoffs defined by equation 7.22. Table 1 took about two hours to produce on a UNIX terminal, so clearly calibrating with the numerical integration to market prices would be unfeasible.

We are now ready to proceed with the calibration. Before we present the results of the calibration we review our data.

7.6 DATA

Our data set was kindly provided by Martin Cooper of Tokai Bank Europe, London. The data consists of contemporaneous money market rates and 2, 3, 5, 7 and 10 year swap rates, for a variety of currencies, together with quoted Black (1976) volatilities for at-the-money caps and swaptions for much of the period 21st June 1995 - 31st May 1996. In more detail, the data set provide Black (1976) volatilities for 1, 2, 3, 4, 5, 7 and 10 years caps, and Black (1976) volatilities for swaptions with maturities 3 months, 6 months, 1, 2, 3, 4, and 5 years on 1, 2, 3, 4, 5, 7 and 10 years swaps. We only use the USD data.

7.7 CALIBRATION TO CAPS AND SWAPTIONS PRICES

We fit to caps and swaptions prices by optimising $g(u, v)$ at the points $(u, v) = \{0,2,4,6,8,10\}^2$ using standard optimisation software to minimise the residual sum of squared differences between model and market prices. The cap and approximation swaption pricing formulae of Section 4 depend on triple integrals of the form

$$\int_{s}^{s_i}\int_{s}^{s_j} c(s,x,y)dydx = \int_{0}^{s}\int_{s-u}^{s_i-u}\int_{s-u}^{s_j-u} g(x,y)dydxdu\,. \tag{7.30}$$

We can obtain an approximation to the integral rapidly using the following procedure. Our choice of caps and instruments imply that we only need to evaluate the integral for s, s_i and s_j that are multiples of a quarter. We interpolate $g(u, v)$ on $\{0, 2, 4, 6, 8, 10\}^2$ to provide values to $\{0, ¼, ½, ¾, 1, 1¼, \ldots, 9¾, 10\}^2$. This allows us to evaluate

$$I_s^{s_i,s_j}(u) = \int_{s-u}^{s_i-u}\int_{s-u}^{s_j-u} g(x,y)dydx$$

analytically when s, s_i, s_j and u are multiples of a quarter. The triple integral of equation 7.30 is approximated using the trapezium rule by evaluating $I_s^{s_i,s_j}(u)$ at u = 0, ¼, ½, ¾, ... , s - ¼, s. With this simplification for the triple integral, we find that the optimisation can be completed in about fifteen minutes on a 75Mhz Pentium PC.

We fit the covariance function simultaneously to market caps and swaptions prices for all dates in May 1996 for which we have data for: 1, 9, 10, 13, 14, 15, 16, 17 20, 21, 22, 23, 28, 29, 30 and 31. There were two national holidays in the UK in May 1996 and the missing dates are clustered around them. We fit to 2, 3, 4, 5, 7 and 10 year caps and 2x1, 2x2, 2x3, 2x4, 2x5, 2x7, 3x1, 3x2, 3x3, 3x4, 3x5, 3x7, 4x1, 4x2, 4x3, 4x4, 4x5, 5x1, 5x2, 5x3, 5x4 and 5x5 swaptions.

The fitted cap and swaptions prices for 31st May 1996 are shown in Tables 2 and 3. The price differences are small. Table 4 shows the Black volatilities for the Kennedy swaptions prices. Except for swaptions on one-year swaps, the fitted Kennedy prices give Black volatilities that are within bid-ask spreads of ± ¼ %. We minimised an equally weighted residual sum of squared pricing differences and because the swaptions on one-year swaps are worth little, their percentage errors are larger and so some Black volatilities are outside the bid-ask range. This can be

easily remedied by adding extra weights to low valued swaptions. We can conclude the fit on 31st May is good. The fit for the other days of May 1996 are similarly good. Figure 1 shows how the residual sum of squares varied.

The fitted function $g(u, v)$ for 31st May 1996 is plotted in Figure 2 and tabulated in Table 5. We find that the fitted functions $g(u, v)$ maintain similar shapes throughout May 1996. Since $g(u, v)$ is the instantaneous covariance of changes to instantaneous forward rates of maturities u and v, some fluctuation through time would be expected. Figure 3 plots the square root of $g(u, u)$, $u = 0..10$ throughout May 1996. The figure shows the volatility term structures of instantaneous forward rates across May 1996. The forward rate volatility term structure maintain a similar shape through time and often exhibit a humped structure often observed in the market. Rather than plotting $g(u, v)$ throughout May 1996, which would be difficult to compare, we examine the implied zero coupon yield covariance structure instead. We construct the matrix C defined by equation 7.13

$$C_{ij} = Cov\left[dY_{0,\tau_i}, dY_{0,\tau_j} \middle| \Im_0\right] = \frac{1}{\tau_i \tau_j} \int_0^{\tau_i} \int_0^{\tau_j} g(u,v)\,dv\,du \qquad \textbf{(7.13)}$$

for τ_i, τ_j = ¼, ½, ¾, ... , 9¾, 10 and extract the eigenvalues and eigenvectors for each day in May 1996. The first three eigenvectors are plotted in Figure 4 and the fourth to the sixth are plotted in Figure 5. All six eigenvectors maintain similar shapes throughout May 1996. This is in marked contrast to methods that extract the volatility factors from historical estimated covariance matrices which produce very unstable higher order eigenvectors. Table 6 shows that for 31st May 1996, the 1st factor account for 95.8% of the variation, the 2nd factor account for 3.2% and the 3rd account for only 0.8%. The first three factors affect zero coupon yields in the precisely the same way that the factors produced by a more traditional PCA decomposition of a historically estimated covariance matrix affect zero coupon

yields. The first factor corresponds to a change in level; the second factor corresponds to a change of slope and the third corresponds to a change of curvature. Figure 6 shows how the first three eigenvalues varied in May 1996 and so the proportions of the zero coupon yield covariance structure explained by each of the first three factors are also stable. An HJM approximation taking just the first three factors would be sufficiently well calibrated to the cap and swaptions prices. This is illustrated in Tables 7 and 8. Table 7 shows the swaptions prices produced by a three-factor HJM approximation to the fitted Kennedy model. The three factor HJM prices are very close to both the market quotes and Kennedy prices. Thus we have a well calibrated HJM model. Table 8 shows the swaptions prices in terms of Black volatilities. All Black volatilities, except for the swaptions on one year swaps are within bid-ask spreads of ± ¼ %. The swaptions on one year swaps have a larger error for the same reason as explained earlier. Tables 7 and 8 also show prices and Black volatilities produced by an one-factor approximation to the fitted Kennedy. The one-factor HJM approximation does not price the swaptions badly. Black volatility errors are greatest for low valued swaptions for the same reason as explained above. Note that it would be possible to make an one factor Gaussian HJM model fit the prices better than the one-factor HJM approximation here. The optimisation routine optimises the fit for Kennedy and not an one factor HJM model. Indeed, Brace, Marek and Musiela (1995) show that one-factor Gaussian HJM models can be fitted to cap and swaption prices simultaneously. In the context of our approach, a one-factor fit would require us to constrain the Kennedy covariance function to be consistent with an one-factor model. This is in contrast to our aim of letting the covariance matrix be unconstrained. Table nine shows the implied correlation of instantaneous zero coupon yield changes.

7.8 SUMMARY

We have shown how Kennedy can be calibrated rapidly and accurately to quoted caps and swaptions prices using an approximate European Swaption pricing formula that we have shown to be accurate for near-the-money swaptions. Our approximate European Swaption pricing formula applies to other diffusion Gaussian interest rate models and so offer them quicker calibration than otherwise. We have shown how our calibrated Kennedy model can be approximated easily and accurately by a Gaussian multifactor HJM. The 3-factor HJM approximation prices caps and swaptions consistently with the market. It also possesses attractive attributes; the volatility factors are stationary and stable between re-calibration across trading days in May 1996.

Our indirect calibration of HJM using Kennedy as an intermediate step is superior to the conventional methods we reviewed in the introduction. We do not have to assume how many factors drive interest rate dynamics and we do not have to assume functional forms for the volatility factors. Our calibration can be achieved quickly.

The Gaussian assumption can of course be a problem. For derivatives where the non-negativity of interest rates is a key factor, Gaussian HJM should not be used to price them. Practitioners may have to resort to non-negative short rate models such as those of Black, Derman and Toy (1990) and Black and Karasinski (1991). More recently, a number of papers have presented HJM type models that do not allow negative interest rates. We examine them in the next chapter.

7.9 APPENDIX

Appendix 1: Hull White (1990) Covariance Function

Hull and White (1990) assume that $r(t)$ follows the risk-neutral process

$$dr(t) = \left[\phi(t) - a(t)r(t)\right]dt + \sigma(t)dz(t).$$

Using a deterministic time change for the Brownian motion $Z(t)$, it is possible to show that $r(t)$ has distribution

$$r(t) \sim c(t) + b(t)\tilde{Z}_{f(t)} \tag{A1}$$

where $\tilde{Z}_{f(t)}$ is another Brownian motion and

$$b(t) = e^{-\int_0^t a(u)du} \tag{A2}$$

$$c(t) = e^{-\int_0^t a(u)du} \int_0^t \phi(u) e^{-\int_0^u a(s)ds} du + r(0) e^{-\int_0^t a(u)du} \tag{A3}$$

$$f(t) = \int_0^t \left[\frac{\sigma(u)}{b(u)} \right]^2 du \tag{A4}$$

Hull and White show that pure discount bond prices are given by

$$P(t,T) = A(t,T)\exp(-B(t,T)r(t)) \tag{A5}$$

where

$$B(t,T) = \frac{B(0,T) - B(0,t)}{\partial B(0,t)/\partial t} \tag{A6}$$

$$B(0,t) = \frac{R(0,t)\sigma_R(0,t)t}{\sigma(0)} \tag{A7}$$

$$\ln A(t,T) = \ln \frac{A(0,T)}{A(0,t)} - B(t,T)\frac{\partial}{\partial t}\ln A(0,t) - \frac{1}{2}\left[B(t,T)\frac{\partial B(0,t)}{\partial t} \right]^2 \int_0^t \left[\frac{\sigma(\tau)}{\partial B(0,\tau)/\partial \tau} \right]^2 d\tau \tag{A8}$$

Equations (A5) and (A6) imply

$$Cov[F(s_1,t_1), F(s_2,t_2)] = \frac{\partial B(0,t_1)/\partial t_1}{\partial B(0,s_1)/\partial s_1} \frac{\partial B(0,t_2)/\partial t_2}{\partial B(0,s_2)/\partial s_2} Cov[r(s_1), r(s_2)]. \tag{A9}$$

Hull and White (1990) show

$$a(t) = -\frac{\partial^2 B(0,t)/\partial t^2}{\partial B(0,t)/\partial t} \tag{A10}$$

which gives by substituting equation (A10) into (A2)

$$b(t) = \exp \int_0^t \frac{\partial^2 B(0,u)/\partial u^2}{\partial B(0,u)/\partial u} du = \frac{\partial B(0,t)}{\partial t}. \tag{A11}$$

Equation (A1) gives

$$Cov[r(s_1), r(s_2)] = b(s_1)b(s_2)Cov[\tilde{Z}_{f(s_1)}, \tilde{Z}_{f(s_2)}] = b(s_1)b(s_2)f(s_1 \wedge s_2) \tag{A12}$$

and substituting equations (A11), (A12) into equation (A9) finally gives

$$Cov[f(s_1,t_1), f(s_2,t_2)] = \frac{\partial B(0,t_1)}{\partial t_1}\frac{\partial B(0,t_2)}{\partial t_2}\int_0^{s_1 \wedge s_2}\left[\frac{\sigma(u)}{\partial B(0,u)/\partial u}\right]^2 du.$$

Appendix 2: Continuity of the Gaussian Random Field

We have from Adler (1981) Theorem 3.4.1 that our real, zero-mean Gaussian Random Field with a continuous covariance function has, with probability one, continuous sample functions over I_0, if there exists some $0 < C < \infty$ and some $\varepsilon > 0$ such that

$$E\left|X(s_1,t_1) - X(s_2,t_2)\right|^2 \le \frac{C}{\left|\log\sqrt{(s_2-s_1)^2 + (t_2-t_1)^2}\right|^{1+\varepsilon}}$$

where

$$I_0 = \{s_1, s_2, t_1, t_2: (s_2 - s_1)^2 + (t_2 - t_1)^2 < 1,\ 0 \le s_1 \le t_1,\ 0 \le s_2 \le t_2\}.$$

We first proof that when $Cov(F(s_1,t_1), F(s_2,t_2)) = c(s_1 \wedge s_2, t_1, t_2) = \int_0^{s_1 \wedge s_2} g(t_1 - u, t_2 - u)du$, $0 \le s_1 \le t_1 < 1$, $0 \le s_2 \le t_2 < 1$, of Assumption 1 the Gaussian Random Field has continuous sample functions, with probability one, when $g(t_1-u, t_2-u)$ is continuous and bounded.

We have by assumption $c(s,t_1,t_2) = \int_0^s g(t_1 - u, t_2 - u)du$. Since $Var[dF(s,t)|\Im_s] = Var[dF(s,s+\tau)|\Im_s] = g(\tau,\tau)$, this implies that $g(\tau,\tau) \ge 0$. Suppose $t_2 \ge t_1$ and $s_2 \ge s_1$. We have that

$$Var[F(s_1,t_1)-F(s_2,t_2)]$$

$$= c(s_1,t_1,t_1)-2c(s_1,t_1,t_2)+c(s_2,t_2,t_2)$$

$$= \int_0^{s_1} g(t_1-u,t_1-u)du - 2\int_0^{s_1} g(t_1-u,t_2-u)du + \int_0^{s_2} g(t_2-u,t_2-u)du$$

$$= \int_0^{s_1} g(t_1-u,t_1-u)du - 2\int_0^{s_1} g(t_1-u,t_2-u)du + (1+k(s_1,s_2,t_2)(s_2-s_1))\int_0^{s_1} g(t_2-u,t_2-u)du$$

for some $\infty > k(s_1, s_2, t_2) \geq 0$ since $g(t_2-u, t_2-u) \geq 0$ and bounded by assumption

$$= k(s_1,s_2,t_2)(s_2-s_1)\int_0^{s_1} g(t_2-u,t_2-u)du +$$

$$\int_0^{s_1} g(t_1-u,t_1-u)+g(t_2-u,t_2-u)-2g(t_1-u,t_2-u)du$$

$$\leq k(s_1,s_2,t_2)(s_2-s_1)\int_0^{s_1} g(t_2-u,t_2-u)du.$$

$$+2\int_0^{s_1}\left[g(t_1-u,t_1-u)\vee g(t_2-u,t_2-u)\right]-g(t_1-u,t_2-u)du.$$

Now note firstly that the non-negative semi-definiteness of $g(u, v)$, implies

$$g(t_1-u,t_1-u)g(t_2-u,t_2-u) \geq g(t_1-u,t_2-u)^2$$

$$g(t_1-u,t_1-u)\vee g(t_2-u,t_2-u) \geq g(t_1-u,t_2-u)$$

which implies

$$g(t_1-u,t_1-u)\vee g(t_2-u,t_2-u)-g(t_1-u,t_2-u)$$

$$= g(t_1-u,t_1-u)\vee g(t_2-u,t_2-u)\frac{M(u,t_1,t_2)}{2}(t_2-t_1)$$

for some $\infty > M(u, t_1, t_2) \geq 0$ since $g(u, v)$ is continuous by assumption.

Thus

$$Var[F(s_1,t_1)-F(s_2,t_2)]$$

$$\leq k(s_1,s_2,t_2)(s_2-s_1)\int_0^{s_1} g(t_2-u,t_2-u)du$$

$$+\int_0^{s_1} g(t_1-u,t_1-u)\vee g(t_2-u,t_2-u)M(u,t_1,t_2)(t_2-t_1)du$$

$$\leq k(s_1,s_2,t_2)(s_2-s_1)\int_0^{s_1} g(t_2-u,t_2-u)du$$

$$+N(s_1,t_1,t_2)(t_2-t_1)\int_0^{s_1} g(t_1-u,t_1-u)\vee g(t_2-u,t_2-u)du$$

where $N(s_1,t_1,t_2)=\underset{u\leq s_1}{Max}\, M(u,t_1,t_2)$

$$\leq \left\{k(s_1,s_2,t_2)(s_2-s_1)+N(s_1,t_1,t_2)(t_2-t_1)\right\}\int_0^{s_1} g(t_1-u,t_1-u)\vee g(t_2-u,t_2-u)du$$

$\leq P(s_1,s_2,t_1,t_2)\left[(s_2-s_1)+(t_2-t_1)\right]$ where $0 \leq P(s_1, s_2, t_1, t_2) < \infty$ and

$$P(s_1,s_2,t_1,t_2)=\left\{k(s_1,s_2,t_2)\vee N(s_1,t_1,t_2)\right\}\int_0^{s_1} g(t_1-u,t_1-u)\vee g(t_2-u,t_2-u)du$$

$$\leq P(s_1,s_2,t_1,t_2)(rCos\theta+rSin\theta)$$

$\leq Qr$ where $0 \leq Q < \infty$ and $Q=\underset{s_1,s_2,t_1,t_2}{Max}\, P(s_1,s_2,t_1,t_2)$.

Note that $r=\sqrt{(s_2-s_1)^2+(t_2-t_1)^2} < 1$, so we only need to show further that

$$Qr \leq \frac{C}{|\log r|^{1+\varepsilon}} \qquad \text{(CC)}$$

for some $0 < C < \infty$ and some $\varepsilon > 0$ for all $0 \leq r < 1$. Now

$$\underset{0\leq r<1}{\text{Min}}\left[\frac{d}{dr}\frac{C}{|\log r|^{1+\varepsilon}}\right]=\frac{C(1+\varepsilon)^{2+\varepsilon}}{(2+\varepsilon)^{2+\varepsilon}}>0$$

so there exists some $0 < C < \infty$ such that

$$\frac{C(1+\varepsilon)^{2+\varepsilon}}{(2+\varepsilon)^{2+\varepsilon}}>Q$$

and so the proposition follows from observing that (CC) holds with equality at $r = 0$ and with strict inequality for $0 < r < 1$. Exchanging t_2 for t_1 and or s_2 for s_1 and adjusting the definition of r and or θ appropriately leads to the same conclusion. To extend the range over which the Proposition applies, we only need to find some T, $\infty > T > max(t_1, t_2)$. Then we can scale time by $1/T$ and the proposition applies. **QED**.

7.10 REFERENCES

1. Adler R.J., 1981, *The Geometry of Random Fields*, (New York: Wiley).
2. Black F., March 1976, "The Pricing of Commodity Contracts", *Journal of Financial Economics* 3.
3. Black F., Derman E. and Toy W., 1990, "A One-Factor Model of Interest Rates and its Application to Treasury Bond Options", *Financial Analysts Journal*, January-February 1990, pp. 33-39.
4. Black F. and Karasinski P., 1991, "Bond and Option Pricing when Short Rates are Lognormal", *Financial Analysts Journal*, July-August 1991, pp. 52-59.
5. Brace A., Gatarek D. and Musiela M., 1995, "The Market Model of Interest Rate Dynamics", *Working Paper, USNW.*
6. Carverhill A. and Pang K., September 1995, "Efficient and Flexible Bond Option Valuation in the Heath, Jarrow and Morton Framework", *Journal of Fixed Income*, Vol. 5 No 2, pp. 70-77.
7. Geman H., El Karoui N., Rochet J-C., 1995, "Changes of Numéraire, Changes of Probability Measure and Option Pricing", *Journal of Applied Probability*, Vol. 32, pp. 443-458.
8. Harrison J.M. and Kreps D.M., 1979, "Martingales and Arbitrage in Multiperiod Securities Markets", *Journal of Economic Theory*, Vol. 20, pp. 381-408.

9. Harrison J.M. and Pliska S.R., 1981, "Martingales and Stochastic Integrals in the Theory of Continuous Trading", *Stoch. Proc. Appl.* Vol. 11, pp. 215-260.

10. Heath D.C., Jarrow R.A. and Morton A., 1992, "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation", *Econometrica*, Vol. 60, pp. 77-105.

11. Hull J. and White A., 1990, "Pricing Interest-Rate Derivative Securities", *Review of Financial Studies*, Vol. 3, pp. 573-592.

12. Hull J. and White A., 1993, "One-Factor Interest Rate Models and the Valuation of Interest-Rate Derivative Securities", *Journal of Financial and Quantitative Analysis*, Vol. 28, pp. 235-254.

13. Hull J. and White A., 1994, "Numerical Procedures for Implementing Term Structure Models II: Two-Factor Models", *The Journal of Derivatives*, Winter 1994, pp. 37-48.

14. Jamshidian F., March 1989, "An Exact Bond Option Formula", *Journal of Finance*, Vol XLIV No 1.

15. Joy C., Boyle P.P. and Tan K.S., February 1995, "Quasi Monte Carlo Methods in Numerical Finance", *Working Paper, University of Waterloo.*

16. Kennedy D.P., 1995, "Characterizing and Filtering Gaussian Models of the Term Structure of Interest Rates", *Working Paper, University of Cambridge.*

17. Kennedy D.P., July 1994, "The Term Structure of Interest Rates as a Gaussian Random Field", *Mathematical Finance*, Vol 4 No 3, pp. 247-258.

18. Langetieg T.C., 1980, "A Multivariate Model of the Term Structure", *Journal of Finance*, Vol XXXV No 1, pp. 71-97.

19. Press,W.H., Teukolsky S.A., Vetterling W.T. and Flannery B.P., 1992, *Numerical Recipes in C*, (Cambridge University Press).

20. Rebonato R. and Cooper I., 1996, "The Limitations of Simple Two-Factor Interest Rate Models", *Journal of Financial Engineering*, Vol 5 No 1.

21. Vasicek O., 1977, "An Equilibrium Characterization of the Term Structure", *J. Financial Econ.*, 5, pp. 177-188.

Figure 1: Residual Sum of Squares

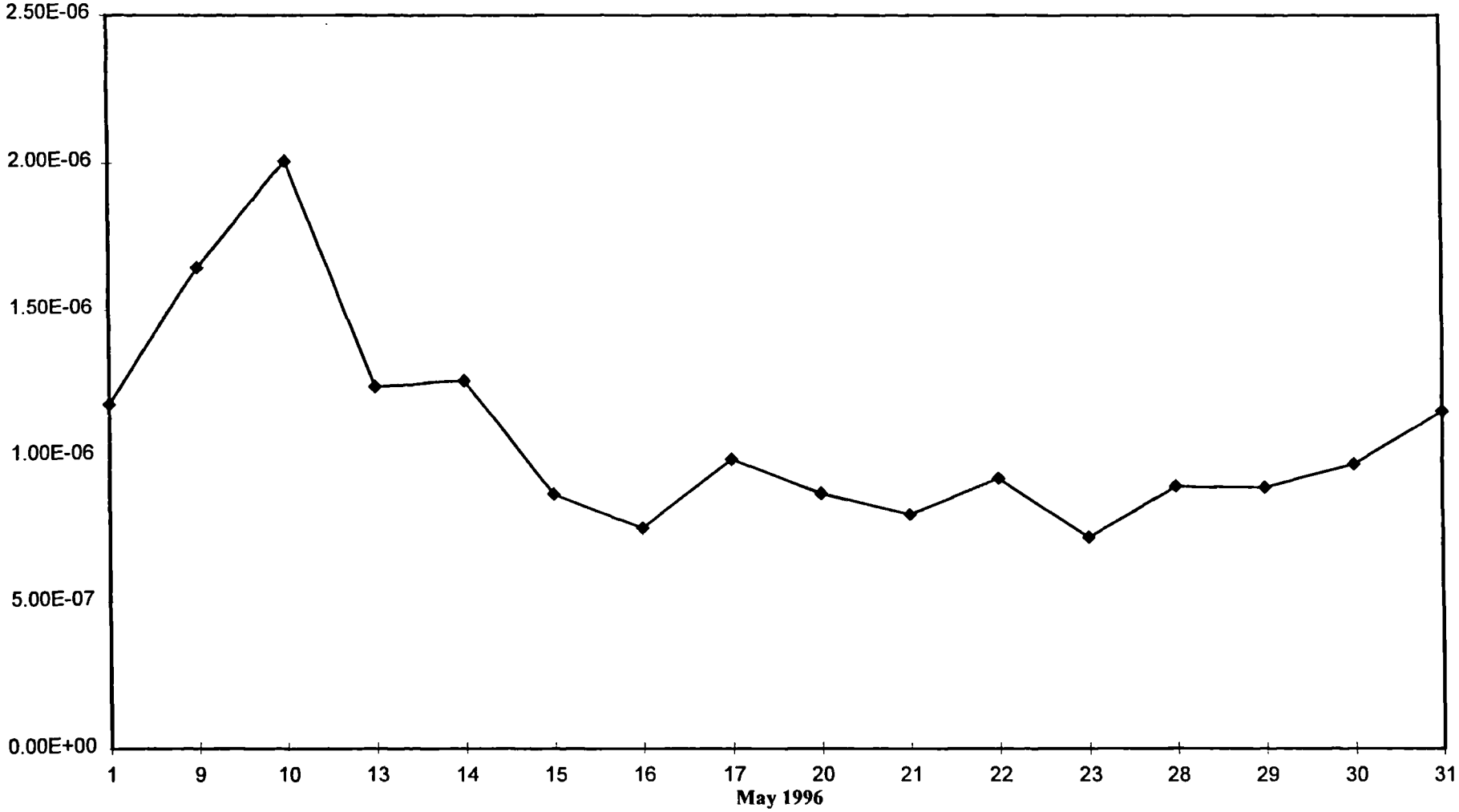


Figure 2: $g(u, v)$, 31st May 1996.

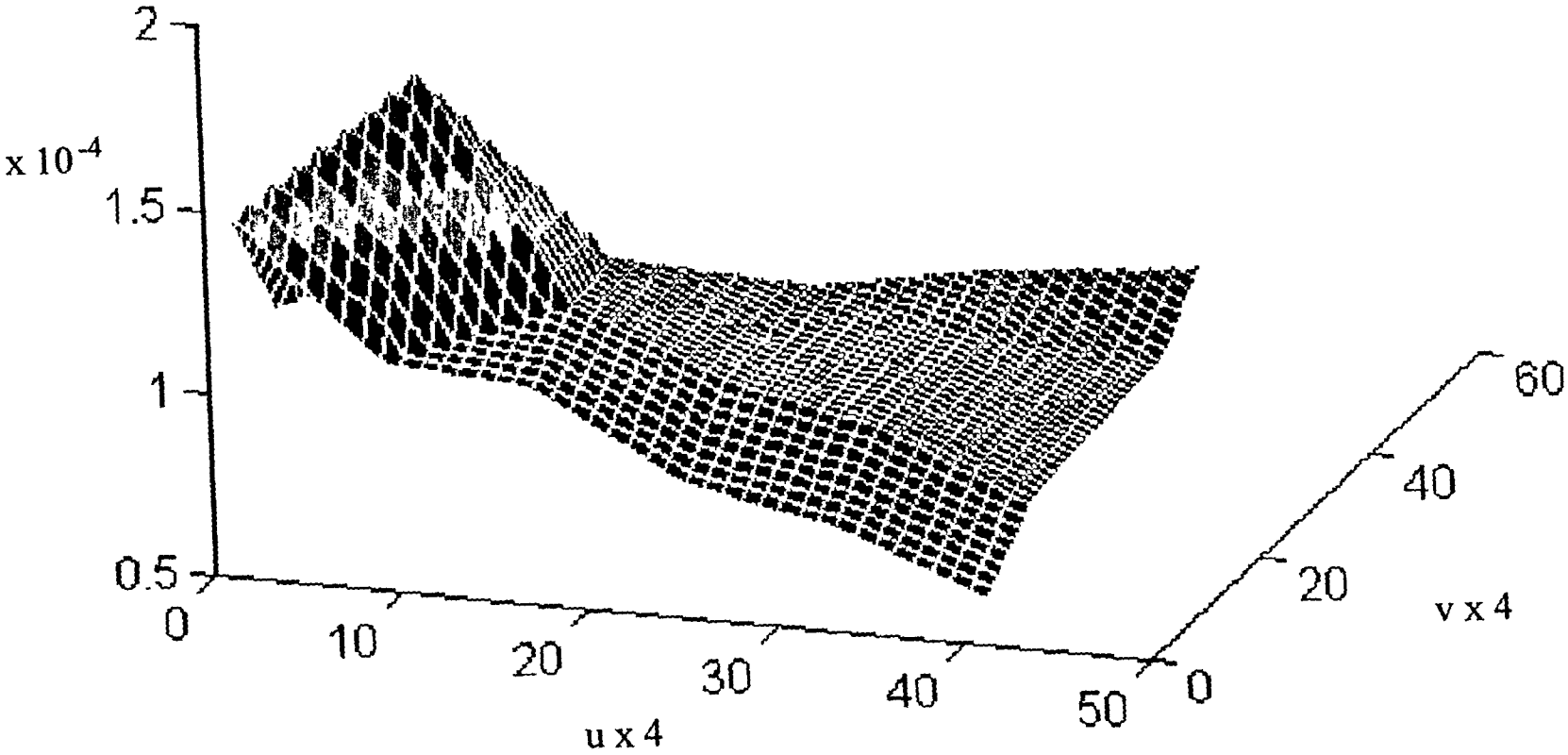


Figure 3: Implied Instantaneous Forward Rate Volatilities
1, 9, 10, 13, 14, 15, 16, 17, 20, 21, 22, 23, 28, 29, 30, 31 May 1996.

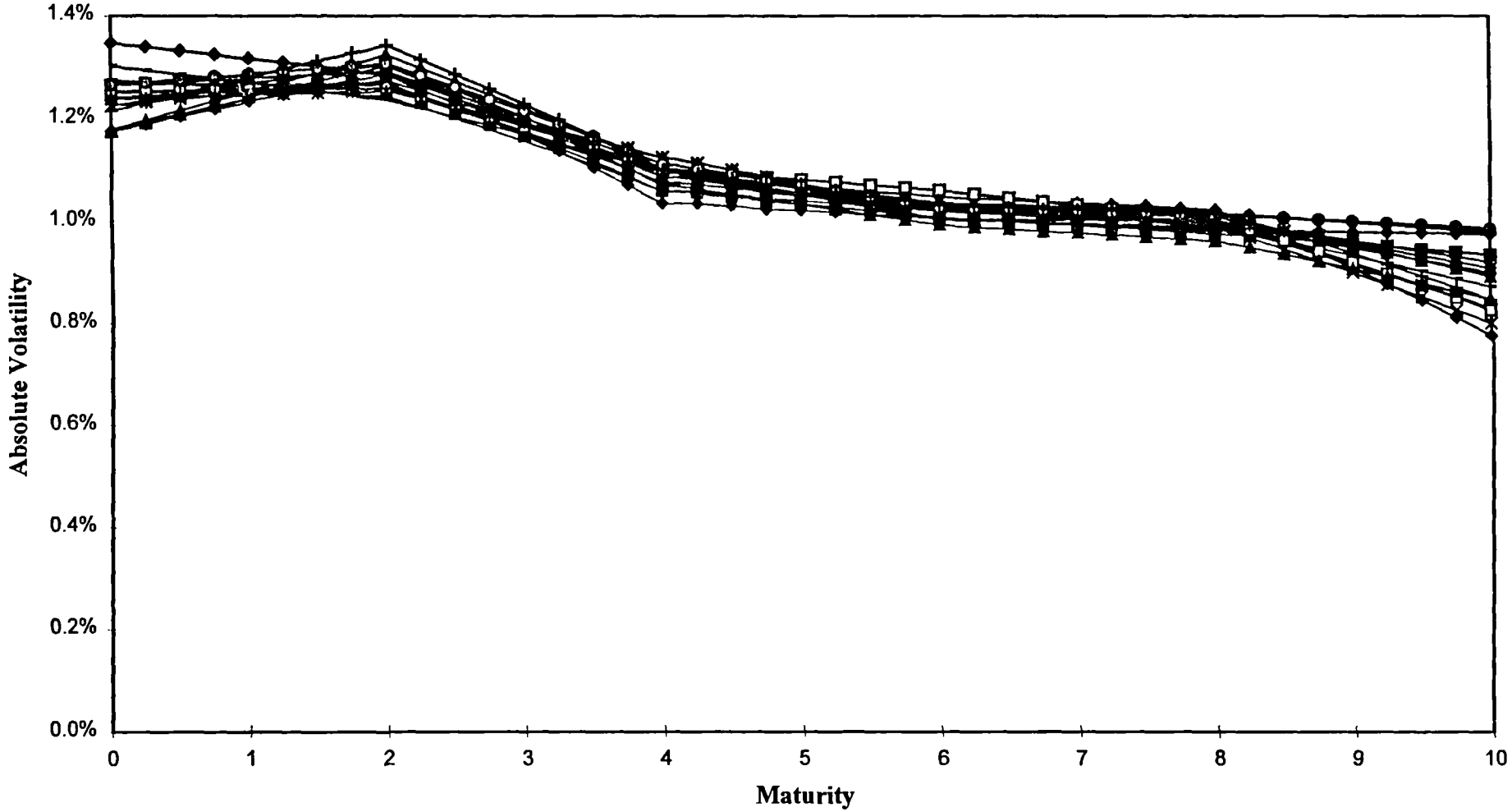
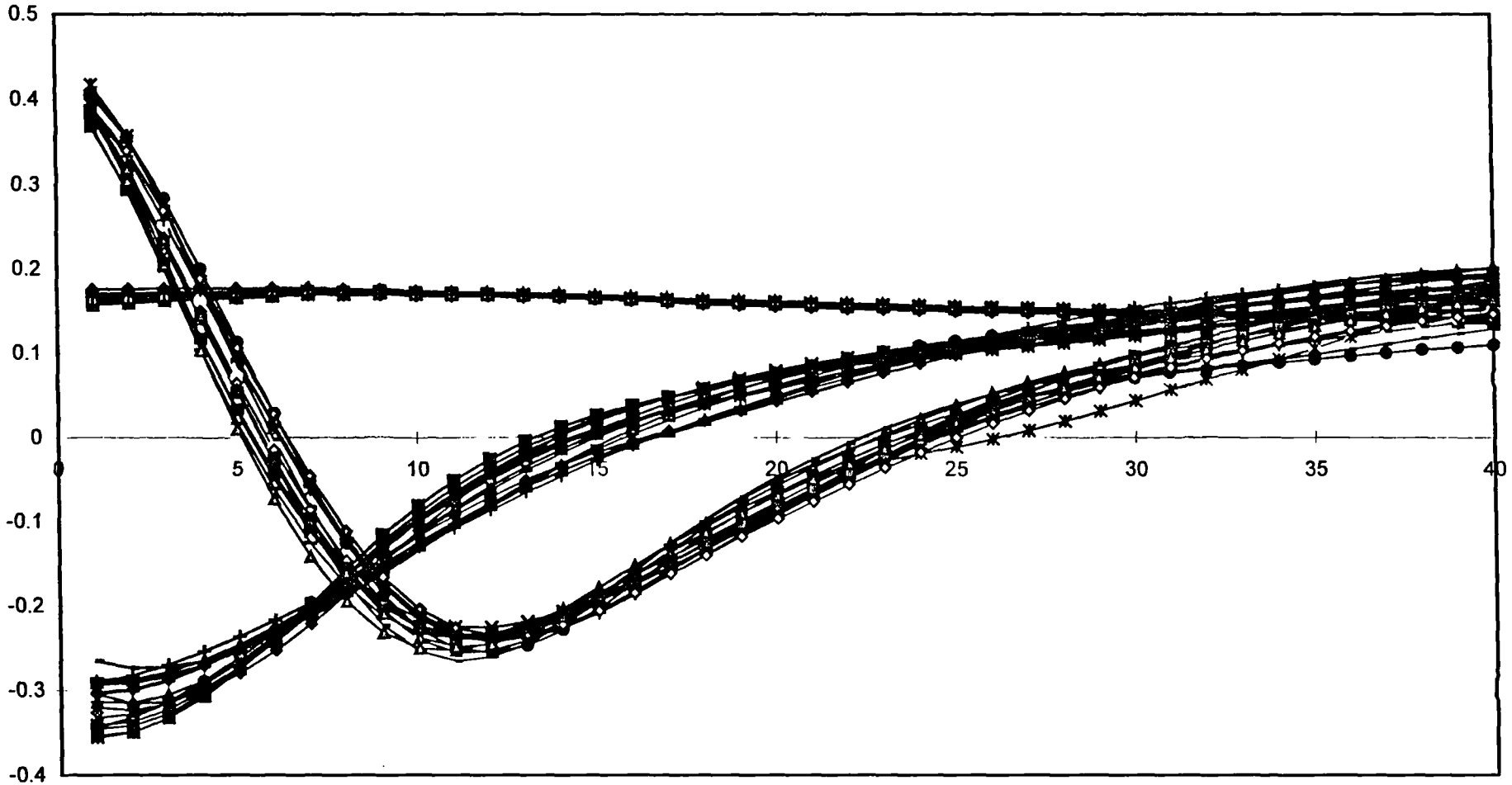
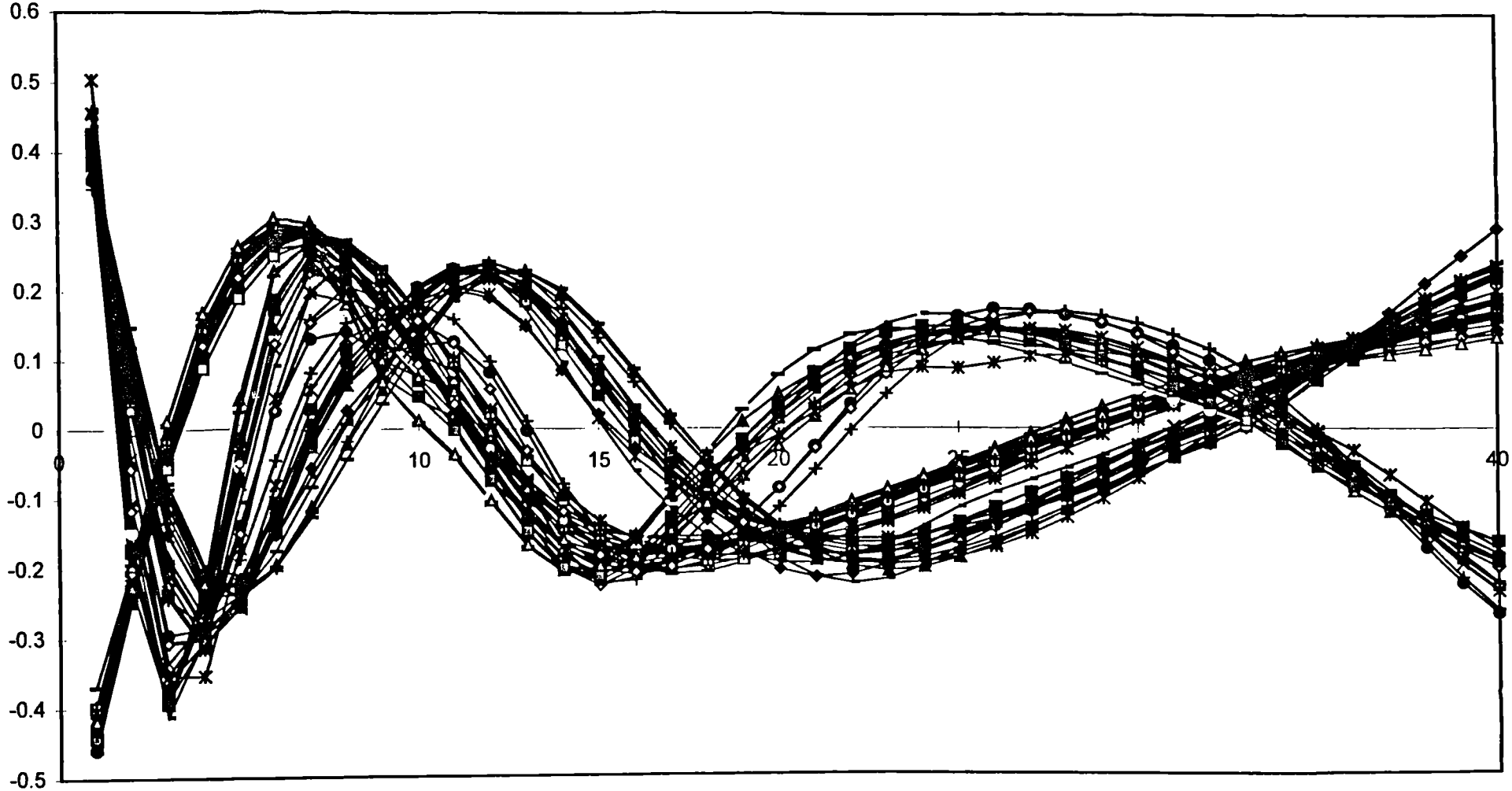


Figure 4: Implied Zero Coupon Yield Volatility Factors, 1st - 3rd
1, 9, 10, 13, 14, 15, 16, 17, 20, 21, 22, 23, 28, 29, 30, 31 May 1996.



Zero Rate Maturity x 4

Figure 5: Implied Zero Coupon Yield Volatility Factors, 4th - 6th
1, 9, 10, 13, 14, 15, 16, 17, 20, 21, 22, 23, 28, 29, 30, 31 May 1996.



Zero Rate Maturity x 4

Figure 6: Principal Eigenvalues

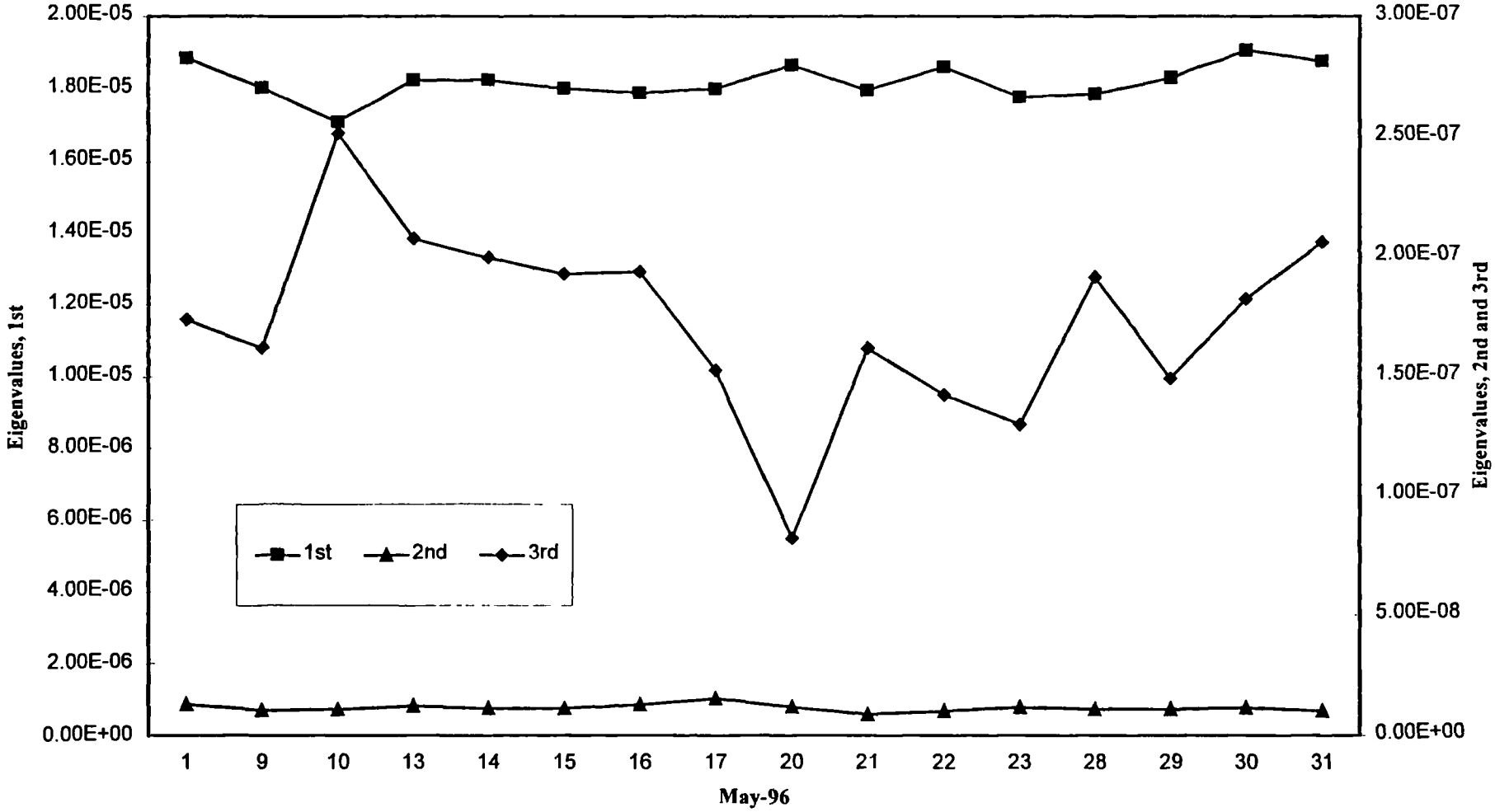


Table 1: Comparison of Numerical Integration andApproximation Formula for European Payer Swaption Values, 31st May 1996.

		Swap Tenor					
		1	2	3	4	5	7
Moneyness[1] 0.9							
2 Year	Num. Integration[2]	0.009644	0.018358	0.026098	0.033044	0.039402	0.050541
Swaptions	Approximation	0.009643	0.01836	0.026112	0.033059	0.039432	0.050592
3 Year	Num. Integration	0.010197	0.019305	0.027318	0.034554	0.041128	0.052601
Swaptions	Approximation	0.010196	0.019313	0.027322	0.034564	0.041134	0.052635
4 Year	Num. Integration	0.010400	0.019576	0.027717	0.035031	0.041689	
Swaptions	Approximation	0.010399	0.019571	0.027714	0.035018	0.041678	
5 Year	Num. Integration	0.010277	0.019414	0.027492	0.03478	0.041377	
Swaptions	Approximation	0.010276	0.019417	0.027489	0.034777	0.041384	
Moneyness 1.0							
2 Year	Num. Integration	0.006305	0.011822	0.016423	0.020421	0.023955	0.029973
Swaptions	Approximation	0.006304	0.011824	0.016433	0.020434	0.023981	0.030031
3 Year	Num. Integration	0.007098	0.013176	0.01834	0.022843	0.026875	0.033664
Swaptions	Approximation	0.007097	0.013181	0.018344	0.022855	0.026892	0.033727
4 Year	Num. Integration	0.007433	0.013828	0.019296	0.024124	0.028425	
Swaptions	Approximation	0.007433	0.013825	0.019297	0.024122	0.028435	
5 Year	Num. Integration	0.007544	0.014057	0.019691	0.024659	0.029064	
Swaptions	Approximation	0.007544	0.014060	0.019692	0.024667	0.02909	
Moneyness 1.1							
2 Year	Num. Integration	0.003823	0.007019	0.009436	0.011432	0.013093	0.015791
Swaptions	Approximation	0.003822	0.007021	0.009446	0.01145	0.013131	0.015885
3 Year	Num. Integration	0.004686	0.008473	0.011535	0.014071	0.016293	0.019826
Swaptions	Approximation	0.004686	0.008477	0.011543	0.014092	0.016332	0.019942
4 Year	Num. Integration	0.00508	0.009307	0.012744	0.015704	0.018263	
Swaptions	Approximation	0.005079	0.009306	0.01275	0.015719	0.018303	
5 Year	Num. Integration	0.005335	0.009769	0.013494	0.01668	0.019423	
Swaptions	Approximation	0.005334	0.009772	0.013501	0.016703	0.019477	

1 Moneyness is defined as strike / forward swap rate for the underlying swap tenor.
2 Based on 1 million samples. Error has $O(10^{-6})$.

Table 2: Fitted Caps Prices, 31st May 1996		
Cap Maturity	**Market Cap Price**	**Model Cap Price**
2	0.010331	0.010606
3	0.017984	0.017971
4	0.025852	0.025804
5	0.034419	0.033964
7	0.050143	0.050034
10	0.072310	0.072806

Table 3: Fitted Swaptions Prices, 31st May 1996						
2 Year Swaptions						
Tenor	1	2	3	4	5	7
Market	0.006329	0.011687	0.016512	0.020442	0.024142	0.030126
Kennedy	0.006304	0.011824	0.016433	0.020434	0.023981	0.030031
3 Year Swaptions						
Tenor	1	2	3	4	5	7
Market	0.006958	0.013153	0.018291	0.023031	0.026725	0.034146
Kennedy	0.007097	0.013181	0.018344	0.022855	0.026892	0.033727
4 Year Swaptions						
Tenor	1	2	3	4	5	-
Market	0.007459	0.013744	0.019442	0.024021	0.028558	-
Kennedy	0.007433	0.013825	0.019297	0.024122	0.028435	-
5 Year Swaptions						
Tenor	1	2	3	4	5	-
Market	0.007396	0.013831	0.019451	0.024347	0.029145	-
Kennedy	0.007544	0.014060	0.019692	0.024667	0.029090	-

Table 4: Fitted Swaptions, Black Volatilities, 31st May 1996						
2 Year Swaptions						
Tenor	1	2	3	4	5	7
Market	19.2	18.2	17.5	16.7	16.2	15.3
Kennedy	19.12	18.41	17.42	16.69	16.09	15.25
3 Year Swaptions						
Tenor	1	2	3	4	5	7
Market	18.2	17.5	16.7	16.2	15.5	15
Kennedy	18.57	17.54	16.75	16.08	15.60	14.81
4 Year Swaptions						
Tenor	1	2	3	4	5	-
Market	17.5	16.7	16.2	15.5	15.2	-
Kennedy	17.44	16.80	16.08	15.57	15.13	-
5 Year Swaptions						
Tenor	1	2	3	4	5	-
Market	16.7	16	15.5	15	14.8	-
Kennedy	17.04	16.27	15.69	15.12	14.77	-

	Table 5: Fitted $g(u, v)$x10^{-5}, 31st May 1996					
	0	**2**	**4**	**6**	**8**	**10**
0	14.78					
2	11.05	18.04				
4	10.74	11.93	12.14			
6	8.50	10.24	10.36	10.63		
8	7.66	9.31	9.44	9.56	10.32	
10	6.07	7.58	7.83	7.95	8.07	9.60

	Table 6: Zero Coupon Yield Volatility Factors, 31st May 1996									
	1st	**2nd**	**3rd**	**4th**	**5th**	**6th**	**7th**	**8th**	**9th**	**10th**
Eigenvalue	1.19e-3	3.95e-5	1.03e-5	2.02e-6	5.54e-7	1.70e-7	6.36e-8	3.11e-8	1.39e-8	5.00e-9
%	95.77%	3.17%	0.83%	0.16%	0.04%	0.01%	0.01%	0.00%	0.00%	0.00%
Cum. %	95.8%	98.9%	99.8%	99.9%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
1	0.327	-0.624	0.602	0.306	-0.176	-0.097	-0.082	-0.035	-0.007	0.000
2	0.340	-0.432	-0.122	-0.604	0.406	0.261	0.257	0.135	0.034	-0.002
3	0.340	-0.178	-0.488	-0.215	-0.273	-0.380	-0.494	-0.310	-0.093	0.012
4	0.333	-0.025	-0.392	0.286	-0.408	-0.045	0.387	0.518	0.251	-0.062
5	0.324	0.070	-0.207	0.372	0.042	0.426	0.255	-0.426	-0.490	0.199
6	0.315	0.156	-0.058	0.293	0.360	0.282	-0.352	-0.129	0.515	-0.415
7	0.306	0.227	0.065	0.139	0.399	-0.199	-0.301	0.439	-0.129	0.576
8	0.298	0.282	0.166	-0.039	0.216	-0.493	0.264	0.013	-0.370	-0.553
9	0.290	0.325	0.246	-0.213	-0.124	-0.194	0.333	-0.421	0.481	0.363
10	0.282	0.355	0.304	-0.352	-0.460	0.441	-0.271	0.219	-0.193	-0.119

Table 7: Fitted Swaptions Prices, 31st May 1996						
2 Year Swaptions Prices						
Tenor	1	2	3	4	5	7
Market	0.006329	0.011687	0.016512	0.020442	0.024142	0.030126
Three-Factor HJM approx.	0.00626	0.01182	0.016425	0.020435	0.02398	0.030034
One-Factor HJM approx.	0.005797	0.011348	0.016175	0.020361	0.024024	0.030076
3 Year Swaptions Prices						
Tenor	1	2	3	4	5	7
Market	0.006958	0.013153	0.018291	0.023031	0.026725	0.034146
Three-Factor HJM approx.	0.007048	0.013177	0.018336	0.022856	0.026892	0.03373
One-Factor HJM approx.	0.006549	0.012686	0.018071	0.022775	0.026935	0.033775
4 Year Swaptions Prices						
Tenor	1	2	3	4	5	-
Market	0.007459	0.013744	0.019442	0.024021	0.028558	-
Three-Factor HJM approx.	0.007423	0.01382	0.019297	0.024121	0.028437	
One-Factor HJM approx.	0.007062	0.01359	0.019254	0.024173	0.028465	
5 Year Swaptions Prices						
Tenor	1	2	3	4	5	-
Market	0.007396	0.013831	0.019451	0.024347	0.029145	-
Three-Factor HJM approx.	0.007532	0.014055	0.019692	0.024665	0.029091	
One-Factor HJM approx.	0.007196	0.013835	0.019647	0.024715	0.029118	

Table 8: Fitted Swaptions, Black Volatilities, 31st May 1996						
2 Year Swaptions						
Tenor	1	2	3	4	5	7
Market	19.2	18.2	17.5	16.7	16.2	15.3
Three-Factor HJM approx.	18.99	18.41	17.41	16.69	16.09	15.25
One-Factor HJM approx.	17.58	17.67	17.14	16.63	16.12	15.27
3 Year Swaptions						
Tenor	1	2	3	4	5	7
Market	18.2	17.5	16.7	16.2	15.5	15
Three-Factor HJM approx.	18.44	17.53	16.74	16.08	15.60	14.82
One-Factor HJM approx.	17.12	16.87	16.50	16.02	15.62	14.84
4 Year Swaptions						
Tenor	1	2	3	4	5	-
Market	17.5	16.7	16.2	15.5	15.2	-
Three-Factor HJM approx.	17.41	16.79	16.08	15.57	15.13	
One-Factor HJM approx.	16.56	16.51	16.04	15.60	15.15	
5 Year Swaptions						
Tenor	1	2	3	4	5	-
Market	16.7	16	15.5	15	14.8	-
Three-Factor HJM approx.	17.01	16.26	15.69	15.12	14.77	
One-Factor HJM approx.	16.24	16.00	15.66	15.15	14.79	

	Table 9: Zero Coupon Yield Instantaneous Change Correlations, 31st May 1996.									
	1 yr.	**2yr.**	**3yr.**	**4yr.**	**5yr.**	**6yr.**	**7yr.**	**8yr.**	**9yr.**	**10yr.**
1 yr.	1									
2 yr.	0.970	1								
3 yr.	0.928	0.984	1							
4 yr.	0.914	0.968	0.994	1						
5 yr.	0.908	0.957	0.985	0.997	1					
6 yr.	0.897	0.944	0.973	0.989	0.997	1				
7 yr.	0.883	0.930	0.961	0.979	0.991	0.998	1			
8 yr.	0.870	0.916	0.948	0.969	0.983	0.993	0.998	1		
9 yr.	0.857	0.903	0.935	0.958	0.975	0.986	0.994	0.998	1	
10 yr.	0.844	0.890	0.923	0.947	0.965	0.979	0.988	0.994	0.998	1

8. THE MARKET-LIBOR MODEL

8.1 INTRODUCTION

In the previous chapter we showed how Gaussian HJM models can be calibrated to caps and swaptions prices. However there are many instruments that cannot be priced well under the Gaussian assumption and for which we need to find an appropriate non-negative interest rate model that can also be calibrated easily to market data. We have also argued in Chapter 4 that tree-based non-negative short rate models like Hull and White (1990, 1994), Black, Derman and Toy(1990), Black and Karasinski(1991) and their multivariate extensions can also be unsatisfactory. So we turn to non-negative HJM models.

There are a several non-negative HJM models in the literature. However, most early non-negative HJM models were artificial in the sense that they were constructed from non-negative short-rate models. That is, until Flesaker and Hughston (1996) provided a complete characterisation of the sub-class of HJM models that guarantees positive interest rates, whilst avoiding infinite interest rates. However, the calibration of their model appears to be formidable, so we shall concentrate on the market-Libor model in this chapter.

The market-Libor model of Brace, Gatarek and Musiela (1995) and Musiela and Rutkowski (1995) provides a set[1] of strictly positive forward Libor rates. In the market-Libor model, the forward Libor rates, of a specified period, are lognormally distributed with respect to different equivalent martingale measures that correspond to well chosen numéraires. The lognormality is consistent with the

---

[1] This set of positive forward Libor rates are characterised by their maturities, $t + m\delta$, where $m = 0, \ldots, N$ and N is the number of Libor rates considered and δ is the Libor rate length.

market convention for the pricing for caplets. The market-Libor model appears therefore to be particular attractive to practitioners.

We will illustrate the advantages of the market-Libor model by examining how resettable caps and floors can be priced by the model. We derive an exact lower bound and an approximate upper bound for the prices of resettable caps and floors. We show that in the market-Libor model there is an exact functional relationship between resettable caplet and floorlet prices. We also provide approximate formulae for the prices of resettable caplets and floorlets. The approximation formulae are very simple and can be implemented easily on a spreadsheet. One of the approximations can price resettable caps and floors directly of quotes for standard caps and floors.

We begin by outlining the main results from Musiela and Rutkowski (1996) needed for us to price the resettable caps and floors. Readers are referred to the paper for full details. Section 8.3 provides formulae for the upper and lower bounds for the value of a resettable caplet. Section 8.4 derives the corresponding results for a resettable floorlet. Section 8.5 provides numerical examples and Section 8.6 provides the approximations. Section 8.7 discusses the advantages and disadvantages of the market-Libor model. Section 8.8 summarises.

8.2 RESULTS FROM MUSIELA AND RUTKOWSKI (1996)

We reproduce some results in this section from Musiela and Rutkowski (1996) that we will use to derive the bounds and approximation formulae for the resettable caps and floors. We use the same notation as Musiela and Rutkowski (1996).

Musiela and Rutkowski (1996) assume we are given a d-dimensional Wiener process $\underline{W}^*$ defined on a filtered probability space $(\Omega, (\Im_t)_{t\in[0,T^*]}, \mathbf{P}^*)$ satisfying the usual assumptions. They assume the forward bond return defined by

$$F_B(t,T,T^*) \stackrel{def}{=} \frac{B(t,T)}{B(t,T^*)}, \; t \le T \le T^*$$

follows a strictly-positive continuous martingale under $\mathbf{P}^*$. See Musiela and Rutkowski (1996), BP.1 - BP.3. $F_B(t, T, T^*)$ is the return from time T to T^* available at time t by shorting one $1/B(t, T^*)$ units of the T^* maturity PDB and investing the unit cash in the T maturity PDB. It is also the value of the T maturity PDB using the T^* maturity PDB as the numéraire. Since $F_B(t, T, T^*)$ is a martingale with respect to $\mathbf{P}^*$, it follows from Harrison and Kreps (1979) that there is no-arbitrage in the PDB market.

Using standard results on strictly-positive continuous martingale, it is possible to express the dynamics of $F_B(t, T, T^*)$ by the Itô differential equation

$$dF_B(t,T,T^*) = F_B(t,T,T^*)\underline{\gamma}(t,T,T^*) \cdot d\underline{W}_t^*$$

where $\underline{\gamma}(t, T, T^*)$, $t \in [0, T]$, is a R^d-valued predictable process, integrable with respect to the Wiener process $\underline{W}^*$ and the $\cdot$ denotes the inner product.

Using some foresight, note that forward Libor rates define forward bond returns over different time intervals. So to consider the forward bond return between any two time points closer than T^*, let us define

$$F_B(t,T,U) \stackrel{def}{=} \frac{F_B(t,T,T^*)}{F_B(t,U,T^*)} = \frac{B(t,T)}{B(t,U)}, \; \forall\, t \in [0, T \wedge U].$$

It follows from an application of Itô's Lemma that

$$\begin{aligned} dF_B(t,T,U) &= F_B(t,T,U)\underline{\gamma}(t,T,U) \cdot (d\underline{W}_t^* - \underline{\gamma}(t,U,T^*)dt) \\ &= F_B(t,T,U)\underline{\gamma}(t,T,U) \cdot d\underline{W}_t^U \end{aligned} \tag{8.1}$$

where

$$\underline{\gamma}(t,T,U) = \underline{\gamma}(t,T,T^*) - \underline{\gamma}(t,U,T^*), \; \forall\, t \in [0, T \wedge U]$$

and

$$\underline{W}_t^U = \underline{W}_t^* - \int_0^t \underline{\gamma}(s,U,T^*)ds\,,\ \forall\ t \in [0,\ U], \tag{8.2}$$

by Girsanov's Theorem, is a Wiener process with respect to the probability measure $\mathbf{P}^U \sim \mathbf{P}^*$ defined by the Radon-Nikodym derivative

$$\frac{d\mathbf{P}^U}{d\mathbf{P}^*} = \exp\left[-\tfrac{1}{2}\int_0^U \left|\underline{\gamma}(s,U,T^*)\right|^2 ds + \int_0^U \underline{\gamma}(s,U,T^*)\cdot d\underline{W}_s^*\right]$$

where $|.|$ denotes the norm in R^d. This change of measure corresponds to a change of numéraire from the T^* maturity PDB to the U maturity PDB. Where convenient, we shall use T^*-measure for $\mathbf{P}^*$ and U-measure for $\mathbf{P}^U$. We also use $E^*[\cdot]$ for expectations taken with respect to the T^*-measure and similarly $E^U[\cdot]$ for expectations taken with respect to the U-measure.

There are various Libor rates with different duration in the market. In the market-Libor model, it is necessary to focus on a particular Libor rate. Let the Libor rate duration be δ. Then the forward δ-Libor rate $L(t,\ T)$ at time t, with maturity T, is given implicitly by

$$1 + \delta\, L(t,T) = F_B(t,T,T+\delta)\,.$$

It follows from equation 8.1, substituting $T+\delta$ for U, and using Itô's Lemma, that under the probability $(T+\delta)$-measure, the forward δ-Libor rate follows the process given by

$$dL(t,T) = L(t,T)\frac{1+\delta\, L(t,T)}{\delta\, L(t,T)}\underline{\gamma}(t,T,T+\delta)\cdot d\underline{W}_t^{T+\delta}\,. \tag{8.3}$$

Examining equation 8.3 reveals that the forward δ-Libor rate $L(t,\ T)$ will be lognormally distributed when

$$\underline{\lambda}(t,T) = \frac{1+\delta\, L(t,T)}{\delta\, L(t,T)}\underline{\gamma}(t,T,T+\delta)$$

is deterministic. The class of model obtained where $\underline{\lambda}(t, T)$ is deterministic is known as the market Libor model[2]. We assume henceforth that $\underline{\lambda}(t, T)$ is deterministic and re-write equation 8.3 as

$$\frac{dL(t,T)}{L(t,T)} = \underline{\lambda}(t,T,) \cdot d\underline{W}_t^{T+\delta} \tag{8.4}$$

which can be solved to give the time T spot δ-Libor rate as

$$L(T,T) = L(t,T)\exp\left\{-\tfrac{1}{2}\int_t^T \left|\underline{\lambda}(u,T)\right|^2 du + \int_t^T \underline{\lambda}(u,T)\cdot d\underline{W}_u^*\right\}. \tag{8.5}$$

It is appropriate to provide a few comments here on the market-Libor model before we proceed to see how the simple Libor derivative, a Libor cap, can be priced. We have seen that when $\underline{\lambda}(t, T)$ is deterministic, the forward δ-Libor rate, $L(t, T)$, is lognormally distributed with respect to $(T+\delta)$-measure. $\underline{\lambda}(t, T)$ is the volatility of $L(t, T)$. Note that other maturity forward Libor rates are not lognormally distributed. That is, $L(t, U)$ is not lognormally distributed with respect to the $(T+\delta)$-measure if U is different from T. Nor are other Libor rates with different lengths lognormally distributed. To see this, let $Z(t, T)$ be the $k\delta$ period forward Libor rate from time T to $T+k\delta$ defined implicitly by

$$1+k\,\delta\, Z(t, T) = B(t, T)/B(t, T+k\delta).$$

Then since

$$\prod_{i=0}^{i=k-1}\left[1+\delta L(t,T+i\delta)\right] = \frac{B(t,T)}{B(t,T+k\delta)}$$

it follows that

---

[2] The assumption that $\lambda(t, T)$ is deterministic is arguably unrealistic and therefore inappropriate but it allows analytical tractability.

$$Z(t,T) = \left[\prod_{i=0}^{i=k-1} \left[1 + \delta L(t, T + i\delta) \right] - 1 \right] \Big/ k\delta .$$

$Z(t, T)$ is not lognormally distributed and is difficult to analyse. These two points we have highlighted are the weaknesses of the model in applications where they require either Libor rates that do not have the same length as δ or where they need to examine different maturing forward Libor rates simultaneously.

The attraction offered by the market-Libor is that the lognormality of the forward Libor rates is consistent with the market quoting convention for cap and floors. It is therefore not surprising that the market Libor model Black gives cap and floor pricing formulae that are consistent with the market convention.

First define

$$T^*_{m\delta} = T^* - m\delta .$$

This notation is convenient because $L(t, T)$ is lognormal in the $(T+\delta)$-measure and so $L(t, T^*_\delta)$ is lognormal in the T^*-measure.

Consider the caplet which pays $\delta[L(T^*_\delta, T^*_\delta) - K]^+$ at time T^*. This caplet can be used to limit the interest rate that applies to a floating loan over the period $[T^*_\delta, T^*]$ to K, the cap rate. It follows from the caplet's time T^* payoff that the time t price of the caplet is given by

$$cpt(t) = \delta P(t, T^*) E^*[L(T^*_\delta, T^*_\delta) - K | \Im_t]^+ . \qquad \textbf{(8.6)}$$

It is a standard result that if equation 8.5, with $T \equiv T^*$, is substituted into equation 8.6 and the expectation taken then we obtain

$$cpt(t) = \delta P(t, T^*) \left\{ L(t, T^*_\delta) N[h(t, T^*_\delta)] - KN[h(t, T^*_\delta) - \varsigma(t, T^*_\delta)] \right\} \qquad \textbf{(8.7)}$$

where

$$h(t, T^*_\delta) = \left[\log \frac{L(t, T^*_\delta)}{K} + \tfrac{1}{2}\varsigma^2(t, T^*_\delta) \right] \Big/ \varsigma(t, T^*_\delta) \qquad \textbf{(8.8)}$$

$$\varsigma^2(t,T_\delta^*) = \int_t^{T_\delta^*} \left|\underline{\lambda}(u,T_\delta^*)\right|^2 du \tag{8.9}$$

and $N[\cdot]$ is the normal distribution function. This is the same as the market convention when equation 8.9 is replaced by

$$\varsigma^2(t,T_\delta^*) = \sigma(T_\delta^*)^2 (T_\delta^* - t) \tag{8.10}$$

where $\sigma(T_\delta^*)$ is the forward-forward volatility[3] of the caplet that covers the period $[T_\delta^*, T]$.

The caplet pricing formula shows that if we can extract the forward-forward volatility of the caplet from quoted market Black volatilities, then equations 8.9 and 8.10 provide a constraint on the forward Libor volatility factors of the market-Libor model. We discuss this observation in more detail in Section 8.7. We will derive in Section 8.6 an approximation that will use only these forward-forward volatilities and the current forward Libor rates to price resettable caps and floors.

We now have sufficient material to price resettable caps and floors using the market-Libor model.

8.3 PRICING OF RESETTABLE CAPLET

A resettable caplet is a standard caplet with the modification that the cap rate is given by the spot δ-Libor rate at a time δ before the caplet maturity. For example, consider the resettable caplet which has at time T_δ^* the payoff given by

---

[3] Cap volatilities are quoted on the understanding that the price of a cap is retrieved by using the quoted cap volatility for all the constituent caplets. Therefore, with the market convention, caplets covering the same period in different caps may be priced using different volatilities. In no-arbitrage models, caplets covering the same period should be priced using the same volatility, the forward-forward volatility. We assume the forward-forward volatilities have been extracted from cap quotes.

$\delta[L(T_\delta^*,T_\delta^*)-L(T_{2\delta}^*,T_{2\delta}^*)]^+$. The payoff at time T_δ^* depends on the difference between the spot δ-Libor rates at times T_δ^* and $T_{2\delta}^*$. This resettable caplet has time t value given by

$$rcpt(t) = P(t,T_\delta^*)E^{T_\delta^*}\delta[L(T_\delta^*,T_\delta^*)-L(T_{2\delta}^*,T_{2\delta}^*)|\Im_t]^+ . \qquad \textbf{(8.11)}$$

We need to solve for $L(T_\delta^*,T_\delta^*)$ and $L(T_{2\delta}^*,T_{2\delta}^*)$ under the T_δ^* - measure to evaluate equation 8.11. We follow Musiela and Rutkowski (1996) and use the notation

$$\varepsilon_T\left(\int_t \underline{\lambda}(u,v)\cdot d\underline{W}_u\right) = \exp\left\{-\tfrac{1}{2}\int_t^T |\underline{\lambda}(u,v)|^2 du + \int_t^T \underline{\lambda}(u,v)\cdot d\underline{W}_u\right\}$$

for the exponential martingale. For example, equation 8.5 can be re-expressed as

$$L(T_\delta^*,T_\delta^*) = L(t,T_\delta^*)\varepsilon_{T_\delta^*}\left(\int_t \underline{\lambda}(u,T_\delta^*)\cdot d\underline{W}_u^*\right). \qquad \textbf{(8.12)}$$

$L(T_{2\delta}^*,T_{2\delta}^*)$ is lognormally under the T_δ^* - measure and so analogous to equation 8.12 we can write

$$L(T_{2\delta}^*,T_{2\delta}^*) = L(t,T_{2\delta}^*)\varepsilon_{T_{2\delta}^*}\left(\int_t \underline{\lambda}(u,T_{2\delta}^*)\cdot d\underline{W}_u^{T_\delta^*}\right). \qquad \textbf{(8.13)}$$

Equation 8.12 needs to be re-written for the T_δ^* - measure. From equation 8.2, with $U \equiv T_\delta^*$, we have

$$\underline{W}_t^{T_\delta^*} = \underline{W}_t^* - \int_0^t \underline{\gamma}(s,T_\delta^*,T^*)ds\,,\ \forall\ t \in [0,\ T_\delta^*] \qquad \textbf{(8.14)}$$

or

$$d\underline{W}_t^{T_\delta^*} = d\underline{W}_t^* - \underline{\gamma}(t,T_\delta^*,T^*)dt\,,\ \forall\ t \in [0,\ T_\delta^*] \qquad \textbf{(8.15)}$$

which when substituted in equation 8.12 and simplified gives

$$L(T_\delta^*,T_\delta^*) = L(t,T_\delta^*)\varepsilon_{T_\delta^*}\left(\int_t \underline{\lambda}(u,T_\delta^*)\cdot d\underline{W}_u^{T_\delta^*}\right)\exp\left[\int_t^{T_\delta^*}\underline{\lambda}(u,T_\delta^*)\cdot\underline{\gamma}(u,T_\delta^*,T^*)du\right]$$

$$= L(t,T_\delta^*)\varepsilon_{T_\delta^*}\left(\int_t \underline{\lambda}(u,T_\delta^*)\cdot d\underline{W}_u^{T_\delta^*}\right)\exp\left[\int_t^{T_\delta^*}\left|\underline{\lambda}(u,T_\delta^*)\right|^2 \frac{\delta L(u,T_\delta^*)}{1+\delta L(u,T_\delta^*)}du\right]. \quad \textbf{(8.16)}$$

We now have the spot Libor rates for equation 8.11 so substituting equations 8.13 and 8.16 into the equation gives

$$rcpt(t) = \delta P(t,T_\delta^*)E^{T_\delta^*}\left[\begin{matrix} L(t,T_\delta^*)\varepsilon_{T_\delta^*}\left(\int_t \underline{\lambda}(u,T_\delta^*)\cdot d\underline{W}_u^{T_\delta^*}\right)\exp\{X\} \\ -L(t,T_{2\delta}^*)\varepsilon_{T_{2\delta}^*}\left(\int_t \underline{\lambda}(u,T_{2\delta}^*)\cdot d\underline{W}_u^{T_\delta^*}\right)\end{matrix}\middle|\Im_t\right]^+ \quad \textbf{(8.17)}$$

where

$$X = \int_t^{T_\delta^*}\left|\underline{\lambda}(u,T_\delta^*)\right|^2 \frac{\delta L(u,T_\delta^*)}{1+\delta L(u,T_\delta^*)}du. \quad \textbf{(8.18)}$$

Unfortunately, the expectation in equation 8.17 cannot be evaluated analytically. We will proceed to derive two simple approximations to 8.17 later. First we will derive upper and lower bounds for equation 8.17.

We deal first with the lower bound. Equation 8.17 is an increasing function of X. Since $X > 0$, we can obtain a lower bound for the value of the resettable caplet by evaluating the expectation analytically at $X = 0$. Using the Lemma of the Appendix we obtain the first proposition.

Proposition 1: *A lower bound value for the time t value of a resettable caplet with payoff* $\delta[L(T_\delta^*,T_\delta^*) - L(T_{2\delta}^*,T_{2\delta}^*)]^+$ *at* T_δ^* *is given by*

$$rcpt(t) > \delta P(t,T_\delta^*)\left[L(t,T_\delta^*)N(d_1) - L(t,T_{2\delta}^*)N(d_2)\right] \quad \textbf{(8.19)}$$

where

$$d_1 = \frac{\log\left(\dfrac{L(t,T_\delta^*)}{L(t,T_{2\delta}^*)}\right)}{\sqrt{v(t,T_{2\delta}^*,T_\delta^*)}} + \frac{1}{2}\sqrt{v(t,T_{2\delta}^*,T_\delta^*)} \quad \textbf{(8.20)}$$

$$d_2 = d_1 - \sqrt{v(t, T_{2\delta}^*, T_\delta^*)} \tag{8.21}$$

$$v(t, T_{2\delta}^*, T_\delta^*) = \int_t^{T_\delta^*} \left|\underline{\lambda}(u, T_\delta^*)\right|^2 du - 2\int_t^{T_{2\delta}^*} \underline{\lambda}(u, T_\delta^*) \cdot \underline{\lambda}(u, T_{2\delta}^*) du + \int_t^{T_{2\delta}^*} \left|\underline{\lambda}(u, T_{2\delta}^*)\right|^2 du. \tag{8.22}$$

Now the upper bound. If follows from equation 8.18 that

$$0 \le X \le \int_t^{T_\delta^*} \left|\underline{\lambda}(u, T_\delta^*)\right|^2 du.$$

Thus we can obtain an upper bound by evaluating equation 8.17 at $X =$ $\int_t^{T_\delta^*} \left|\underline{\lambda}(u, T_\delta^*)\right|^2 du$. However, this upper bound will be much too large by far. We provide an alternative upper bound; it holds only approximately but in practice, as we shall see in Section 8.5, it too is much larger than a numerical integration of equation 8.17.

Proposition 2: *An approximate upper bound for the time t value of a resettable caplet with payoff* $\delta[L(T_\delta^*, T_\delta^*) - L(T_{2\delta}^*, T_{2\delta}^*)]^+$ *at* T_δ^* *is given by*

$$rcpt_U(t) \lesssim \frac{Rcpt_M(t) + Rcpt_E(t)}{2} \tag{8.23}$$

where

$$Rcpt_E(t) = \delta P(t, T_\delta^*) f\left[E^{T_\delta^*}\left[X | \Im_t\right]\right] \tag{8.24}$$

$$Rcpt_M(t) = \delta P(t, T_\delta^*) f\left[\int_t^{T_\delta^*} \left|\underline{\lambda}(u, T_\delta^*)\right|^2 du\right] \tag{8.25}$$

$$f(X) = L(t, T_\delta^*)\exp[X]N(d_1) - L(t, T_{2\delta}^*)N(d_2) \tag{8.26}$$

$$d_1 = \frac{\log\left(\frac{L(t,T_\delta^*)}{L(t,T_{2\delta}^*)}\right) + X}{\sqrt{v(t,T_{2\delta}^*,T_\delta^*)}} + \frac{1}{2}\sqrt{v(t,T_{2\delta}^*,T_\delta^*)} \tag{8.27}$$

$$d_2 = d_1 - \sqrt{v(t,T_{2\delta}^*,T_\delta^*)} \tag{8.28}$$

$$v(t,T_{2\delta}^*,T_\delta^*) = \int_t^{T_\delta^*} \left|\underline{\lambda}(u,T_\delta^*)\right|^2 du - 2\int_t^{T_{2\delta}^*} \underline{\lambda}(u,T_\delta^*)\cdot\underline{\lambda}(u,T_{2\delta}^*)du + \int_t^{T_{2\delta}^*} \left|\underline{\lambda}(u,T_{2\delta}^*)\right|^2 du. \tag{8.29}$$

Proof: We have

$$rcpt(t) = \delta P(t,T_\delta^*)E^{T_\delta^*}\left[\begin{array}{l} L(t,T_\delta^*)\varepsilon_{T_\delta^*}\left(\int_t \underline{\lambda}(u,T_\delta^*)\cdot d\underline{W}_u^{T_\delta^*}\right)\exp[X] \\ -L(t,T_{2\delta}^*)\varepsilon_{T_{2\delta}^*}\left(\int_t \underline{\lambda}(u,T_{2\delta}^*)\cdot d\underline{W}_u^{T_\delta^*}\right) \end{array}\middle|\Im_t\right]^+$$

$$= \delta P(t,T_\delta^*)E^{T_\delta^*}\left[E^{T_\delta^*}\left[\begin{array}{l} L(t,T_\delta^*)\varepsilon_{T_\delta^*}\left(\int_t \underline{\lambda}(u,T_\delta^*)\cdot d\underline{W}_u^{T_\delta^*}\right)\exp[X] \\ -L(t,T_{2\delta}^*)\varepsilon_{T_{2\delta}^*}\left(\int_t \underline{\lambda}(u,T_{2\delta}^*)\cdot d\underline{W}_u^{T_\delta^*}\right) \end{array}\middle|\Im_t, X\right]^+\middle|\Im_t\right]$$

$$= \delta P(t,T_\delta^*)E^{T_\delta^*}\left[f(X)|\Im_t\right] \tag{8.30}$$

where $f(X)$, $d_1(X)$ and $d_2(X)$ are as defined in equations 8.26 to 8.28 and X is defined by equation 8.18. Note that equation $f(X)$ is an increasing function of X in the same way that that the Black-Scholes European Call option formula is increasing with the current share price. Now for any increasing function $f(X)$,

$$E[f(X)] \le [f(\text{med}[X]) + f(\max[X])]/2$$

where med[X] is the median of X. We approximate med(X) by E(X), that is, we assume

$$E[f(X)] \lesssim [f(E(X)) + f(\max[X]]/2. \tag{8.31}$$

Now from the definition of X in equation 8.18, we have

$$E^{T_\delta^*}[X|\Im_t] = \int_t^{T_\delta^*} \left|\underline{\lambda}(u,T_\delta^*)\right|^2 E^{T_\delta^*}\left[\frac{\delta L(u,T_\delta^*)}{1+\delta L(u,T_\delta^*)}\middle|\Im_t\right]du$$

$$= \frac{\delta L(t,T_\delta^*)}{1+\delta L(t,T_\delta^*)} \int_t^{T_\delta^*} \left|\underline{\lambda}(u,T_\delta^*)\right|^2 du \qquad \textbf{(8.32)}$$

since

$$\frac{\delta L(u,T_\delta^*)}{1+\delta L(u,T_\delta^*)} = 1 - \frac{1}{1+\delta L(u,T_\delta^*)} = 1 - \frac{P(u,T^*)}{P(u,T_\delta^*)}$$

is a $P^{T_\delta^*}$ martingale. Furthermore,

$$\max[X] = \int_t^{T_\delta^*} \left|\underline{\lambda}(u,T_\delta^*)\right|^2 du. \qquad \textbf{(8.33)}$$

The approximation follows by substituting equations 8.32 and 8.33 into equation 8.31.

QED.

We have derived an exact lower bound and an approximate upper bound for the price of resettable caplets. We will proceed to examine how well these bounds perform in Section 8.5. Before that, we show how the bounds for resettable floorlets can be obtained from the bounds on the resettable caplets.

8.4 PRICING OF RESETTABLE FLOOR

Resettable floorlet have a similar payoff structure to resettable caplets. The resettable floorlet that covers the same period as the resettable caplet considered in Section 8.3 has payoff at time T_δ^* the quantity given by $\delta[L(T_{2\delta}^*,T_{2\delta}^*) - L(T_\delta^*,T_\delta^*)]^+$. So resettable floorlets are the counterpart to resettable caplets. Resettable floorlets pay out when the spot δ-Libor rate $L(T_{2\delta}^*,T_{2\delta}^*)$ is greater than $L(T_\delta^*,T_\delta^*)$ whereas for resettable caplets it is the other way round.

We can derive bounds for the price of the resettable floorlet by considering a portfolio that consists a long position in the resettable caplet considered in Section 8.3 and one short of the resettable floorlet that covers the same period. This portfolio has payoff at time T_δ^* the quantity given by

$$\delta[L(T_\delta^*,T_\delta^*)-L(T_{2\delta}^*,T_{2\delta}^*)]$$

which has the same value as

$$[1+\delta L(T_\delta^*,T_\delta^*)]\delta L(T_\delta^*,T_\delta^*) \text{ at time } T$$

by rolling the amount $\delta L(T_\delta^*,T_\delta^*)$ from time T_δ^* to T and

$$-\delta L(T_{2\delta}^*,T_{2\delta}^*) \text{ at time } T_\delta^*.$$

We write the payoff in this way because $L(T_\delta^*,T_\delta^*)$ is lognormally distributed with respect to the T-measure and $L(T_{2\delta}^*,T_{2\delta}^*)$ is lognormally distributed with respect to the T_δ^*-measure.

The value of the portfolio must be the value of the risk-adjusted discounted expected values of the two cashflows at times T and T_δ^*. Therefore

$$\begin{aligned} rcpt(t)-rflt(t) = {} & P(t,T^*)E^*\left[[1+\delta L(T_\delta^*,T_\delta^*)]\delta L(T_\delta^*,T_\delta^*)\middle|\Im_t\right] \\ & -P(t,T_\delta^*)E^{T_\delta^*}\left[\delta L(T_{2\delta}^*,T_{2\delta}^*)\middle|\Im_t\right]. \end{aligned} \tag{8.34}$$

We can evaluate equation 8.34 analytically using standard results for the moments of the lognormal distribution equation to give

$$\begin{aligned} rflt(t) = {} & rcpt(t)+\delta P(t,T_\delta^*)L(t,T_{2\delta}^*) \\ & -\delta P(t,T^*)\left[L(t,T_\delta^*)+\delta L^2(t,T_\delta^*)\exp\left(\sigma^2(T_\delta^*)(T_\delta^*-t)\right)\right]. \end{aligned} \tag{8.35}$$

The upper and lower bounds for the resettable floorlet are given by substituting the upper and lower bounds for the resettable caplet into equation 8.35.

8.5 NUMERICAL RESULTS

In this section, we investigate where the no-arbitrage prices of resettable caplets are in relation to be bounds derived in Section 8.3. We assume that δ is ¼.

Empirical data suggest interest rate dynamics can be explained by two or three factors. In our example we take two. The first and second factors are shown in Figure 1. The first factor accounts for 90% of variations in interest rates and the second factor accounts for the remainder. The resulting volatility term structure and our assumed initial term structure of forward Libor rates are plotted in Figure 2. Our volatility factors are proportional and are assumed to be time stationary, that is, they depend only on the maturities of the forward Libor rates. We have chosen the volatility factors to give a humped volatility term structure that occurs generally in practice.

Using these inputs, we can evaluate the integrals of Propositions 1 and 2 numerically to give the lower and upper bounds for the price of resettable caplets. The results are plotted in Figure 3. We see that the spread between the approximate upper and the exact lower bounds is disappointingly large. This, however, would not be important in practice if the exact price is close to the lower bound. We examine this possibility using Monte Carlo simulations.

8.5.1 Monte Carlo Simulations

We repeat equation 8.11 here for the value of the resettable caplet covering the period $[T^*_{2\delta}, T^*_\delta]$

$$rcpt(t) = P(t, T^*_\delta) E^{T^*_\delta} \delta [L(T^*_\delta, T^*_\delta) - L(T^*_{2\delta}, T^*_{2\delta}) | \Im_t]^+ .$$

We need to be able to simulate simultaneously the values of $L(T^*_{2\delta}, T^*_{2\delta})$ and $L(T^*_\delta, T^*_\delta)$ under the T^*_δ-measure. Substituting $T^*_{2\delta}$ and T^*_δ for T in equation 8.4 gives the following processes for $L(t, T^*_{2\delta})$ and $L(t, T^*_\delta)$ respectively:

$$\frac{dL(t,T_{2\delta}^*)}{L(t,T_{2\delta}^*)} = \underline{\lambda}(t,T_{2\delta}^*)\cdot d\underline{W}_t^{T_\delta^*} \tag{8.36}$$

$$\frac{dL(t,T_\delta^*)}{L(t,T_\delta^*)} = \underline{\lambda}(t,T_\delta^*)\cdot d\underline{W}_t^{T^*}. \tag{8.37}$$

We can transform equation 8.37 to the T_δ^*-measure using equation 8.15 to give

$$\frac{dL(t,T_\delta^*)}{L(t,T_\delta^*)} = \left|\underline{\lambda}(t,T_\delta^*)\right|^2\left[\frac{\delta L(t,T_\delta^*)}{1+\delta L(t,T_\delta^*)}\right]dt + \underline{\lambda}(t,T_\delta^*)\cdot d\underline{W}_t^{T_\delta^*}. \tag{8.38}$$

Equations 8.36 and 8.38 can be re-expressed as

$$d\ln L(t,T_{2\delta}^*) = -\tfrac{1}{2}\left|\underline{\lambda}(t,T_{2\delta}^*)\right|^2 dt + \underline{\lambda}(t,T_{2\delta}^*)\cdot d\underline{W}_t^{T_\delta^*} \tag{8.39}$$

$$d\ln L(t,T_\delta^*) = \left|\underline{\lambda}(t,T_\delta^*)\right|^2\left[\frac{\delta L(t,T_\delta^*)}{1+\delta L(t,T_\delta^*)} - \frac{1}{2}\right]dt + \underline{\lambda}(t,T_\delta^*)\cdot d\underline{W}_t^{T_\delta^*}. \tag{8.40}$$

Equations 8.39 and 8.40 allow us to simulate for $L(T_{2\delta}^*, T_{2\delta}^*)$ and $L(T_\delta^*, T_\delta^*)$ jointly. It is better to simulate the logarithms because simulating the logarithms prevents the Libor rates going negative that can occur if we simulate the levels using equations 8.36 and 8.38 instead.

Figure 4 plots the lower bound provided by Proposition 1 and the Monte Carlo estimates together with the two standard error band. For caplets with maturities 0.5 to 5.25 we used step sizes of 0.00125 and for the caplets with maturities 5.5 to 9.5 we used 0.025. For both sets we used 10000 simulations with antithetic variance reduction.

We see, for the Libor curve and volatility factors we have chosen, the Monte Carlo prices are not very close to the exact lower bound. The Monte Carlo prices exceed the exact lower bounds by between 5% and 17%. It is obvious that the upper bound plotted in figure 3 is much too high to be practically useful. The lower bound is also too low so it may be more useful to obtain an approximate

formula for the prices of resettable caplets. We provide two approximations in the next section.

8.6 APPROXIMATIONS

We saw in the previous section that the bounds are disappointingly far away from the Monte Carlo Simulation prices for them to be practically useful. In this section to derive approximations that may be more useful. Here is the first approximation.

Approximation 1: *An approximate value for the time t value of a resettable caplet with payoff* $\delta[L(T_\delta^*,T_\delta^*)-L(T_{2\delta}^*,T_{2\delta}^*)]^+$ *at* T_δ^* *is given by*

$$rcpt(t) \approx \delta P(t,T_\delta^*)\left[L(t,T_\delta^*)\exp\left(\frac{\delta L(t,T_\delta^*)}{1+\delta L(t,T_\delta^*)}\int_t^{T_\delta^*}\left|\underline{\lambda}(u,T_\delta^*)\right|^2 du\right)N(d_1)-L(t,T_{2\delta}^*)N(d_2)\right] \tag{8.41}$$

where

$$d_1 = \frac{\log\left(\frac{L(t,T_\delta^*)}{L(t,T_{2\delta}^*)}\right)+\frac{\delta L(t,T_\delta^*)}{1+\delta L(t,T_\delta^*)}\int_t^{T_\delta^*}\left|\underline{\lambda}(u,T_\delta^*)\right|^2 du}{\sqrt{v(t,T_{2\delta}^*,T_\delta^*)}}+\frac{1}{2}\sqrt{v(t,T_{2\delta}^*,T_\delta^*)} \tag{8.42}$$

$$d_2 = d_1-\sqrt{v(t,T_{2\delta}^*,T_\delta^*)} \tag{8.43}$$

$$v(t,T_{2\delta}^*,T_\delta^*) = \int_t^{T_\delta^*}\left|\underline{\lambda}(u,T_\delta^*)\right|^2 du-2\int_t^{T_{2\delta}^*}\underline{\lambda}(u,T_\delta^*)\cdot\underline{\lambda}(u,T_{2\delta}^*)du+\int_t^{T_{2\delta}^*}\left|\underline{\lambda}(u,T_{2\delta}^*)\right|^2 du. \tag{8.44}$$

Proof: We use the approximation

$$E^{T_\delta^*}\left[L(t,T_\delta^*)\varepsilon_{T_\delta^*}\left(\int_t\underline{\lambda}(u,T_\delta^*)\cdot d\underline{W}_u^{T_\delta^*}\right)\exp[X]-L(t,T_{2\delta}^*)\varepsilon_{T_{2\delta}^*}\left(\int_t\underline{\lambda}(u,T_{2\delta}^*)\cdot d\underline{W}_u^{T_\delta^*}\right)\middle|\Im_t\right]^+$$

$$\approx E^{T_\delta^*}\left[\begin{array}{l} L(t,T_\delta^*)\exp\left\{E^{T_\delta^*}\left(X|\Im_t\right)\right\}\varepsilon_{T_\delta^*}\left(\int_t \underline{\lambda}(u,T_\delta^*)\cdot d\underline{W}_u^{T_\delta^*}\right) \\ -L(t,T_{2\delta}^*)\varepsilon_{T_{2\delta}^*}\left(\int_t \underline{\lambda}(u,T_{2\delta}^*)\cdot d\underline{W}_u^{T_\delta^*}\right) \end{array}\Bigg|\Im_t\right]^+ . \tag{8.45}$$

Thus substituting equation 8.45 and equation 8.32 into equation 8.17 gives

$$rcpt(t) \approx \delta P(t,T_\delta^*)E^{T_\delta^*}\left[\begin{array}{l} L(t,T_\delta^*)\exp\left(\frac{\delta L(t,T_\delta^*)}{1+\delta L(t,T_\delta^*)}\int_t^{T_\delta^*}\left|\underline{\lambda}(u,T_\delta^*)\right|^2 du\right)\varepsilon_{T_\delta^*}\left(\int_t \underline{\lambda}(u,T_\delta^*)\cdot d\underline{W}_u^{T_\delta^*}\right) \\ -L(t,T_{2\delta}^*)\varepsilon_{T_{2\delta}^*}\left(\int_t \underline{\lambda}(u,T_{2\delta}^*)\cdot d\underline{W}_u^{T_\delta^*}\right) \end{array}\Bigg|\Im_t\right]^+ .$$

Finally, applying the Lemma of the Appendix gives the required formula. **QED**.

Like the upper and lower bounds, Approximation 1 can be evaluated easily once the volatility structure of the forward Libor rates are specified. Before we proceed to examine how well Approximation 1 performs, let us first examine its connection to quoted cap volatilities. It would be very convenient for practitioners, if it were possible to price, even approximately, resettable caps and floors[4] directly from quoted standard cap and floor volatilities.

Actually, nearly all the inputs required for the bounds and Approximation 1 can be obtained directly from market quotes at time t, assuming that the forward-forward volatilities for the standard caplets have been extracted quoted cap volatilities.

To see this, suppose $\sigma(T_{2\delta}^*)$ represents the forward-forward volatility for a standard caplet that covers the period $[T_{2\delta}^*, T_\delta^*]$ and similarly $\sigma(T_\delta^*)$ the forward-

[4] Each resettable cap (floor) is made up of a strip of end-to-end resettable caplets (floorlets) that cover the entire duration of the resettable cap (floor). The price of resettable caps (floors) are given by the sum of the constituent resettable caplets (floorlets).

forward volatility for $[T_\delta^*, T]$. The caplet pricing formula given by equations 8.7 to 8.9 gives the following constraints on the forward Libor volatility factors:

$$\int_t^{T_{2\delta}^*} \left|\underline{\lambda}(u, T_{2\delta}^*)\right|^2 du = \sigma^2(T_{2\delta}^*)(T_{2\delta}^* - t) \tag{8.46}$$

$$\int_t^{T_\delta^*} \left|\underline{\lambda}(u, T_\delta^*)\right|^2 du = \sigma^2(T_\delta^*)(T_\delta^* - t)\,. \tag{8.47}$$

These constraints are important in the calibration of the market-Libor model and we will discuss calibration in Section 8.6. So given the forward-forward caplets and the initial forward Libor rate term structure, we only need

$$\int_t^{T_{2\delta}^*} \underline{\lambda}(u, T_\delta^*) \cdot \underline{\lambda}(u, T_{2\delta}^*) du = Y, \text{ say,}$$

to be able to evaluate the bounds and Approximation 1.

However, Y cannot be determined without calibrating the model properly. A calibration will typically make assumptions on the forward Libor volatility factors, $\underline{\lambda}(u, T)$, that are consistent with the constraints provided by equations 8.46 and 8.47.

Sometimes practitioners may just want an indication of what the arbitrage free price is without going through the full calibration. To this end, we provide an approximation for Y that gives Approximation 2 for the prices of resettable caplets.

We are looking for an approximation to Y. Note that

$$\left[\int_t^{T_{2\delta}^*} \underline{\lambda}(u, T_\delta^*) \cdot \underline{\lambda}(u, T_{2\delta}^*) du\right]^2 \le \left[\int_t^{T_{2\delta}^*} \left|\underline{\lambda}(u, T_\delta^*)\right| \cdot \left|\underline{\lambda}(u, T_{2\delta}^*)\right| du\right]^2 \le \int_t^{T_{2\delta}^*} \left|\underline{\lambda}(u, T_\delta^*)\right|^2 du \int_t^{T_{2\delta}^*} \left|\underline{\lambda}(u, T_{2\delta}^*)\right|^2 du\,.$$

We have from the caplet forward-forward volatilities

$$\int_t^{T_{2\delta}^*} \left|\underline{\lambda}(u, T_{2\delta}^*)\right|^2 du = \sigma^2(T_{2\delta}^*)(T_{2\delta}^* - t)$$

but not

$$\int_t^{T_{2\delta}^*} \left|\underline{\lambda}(u,T_\delta^*)\right|^2 du .$$

We can approximate the latter by

$$\int_t^{T_{2\delta}^*} \left|\underline{\lambda}(u,T_\delta^*)\right|^2 du \approx \frac{T_{2\delta}^* - t}{T_\delta^* - t} \int_t^{T_\delta^*} \left|\underline{\lambda}(u,T_\delta^*)\right|^2 du = \sigma^2(T_\delta^*)(T_{2\delta}^* - t) . \tag{8.48}$$

and approximate Y by

$$Y = \int_t^{T_{2\delta}^*} \lambda(u,T_\delta^*) \cdot \lambda(u,T_{2\delta}^*) du \approx \sigma(T_{2\delta}^*)\sigma(T_\delta^*)(T_{2\delta}^* - t) . \tag{8.49}$$

Thus we have

Approximation 2: *An alternative approximation for the time t value of a resettable caplet with payoff* $\delta[L(T_\delta^*,T_\delta^*) - L(T_{2\delta}^*,T_{2\delta}^*)]^+$ *at* T_δ^* *is given by*

$$rcpt_L(t) \approx \delta P(t,T_\delta^*)\left[L(t,T_\delta^*)\exp\left(\frac{\delta L(t,T_\delta^*)}{1+\delta L(t,T_\delta^*)}\sigma^2(T_\delta^*)(T_\delta^* - t)\right)N(d_1) - L(t,T_{2\delta}^*)N(d_2)\right] \tag{8.50}$$

$$d_1 = \frac{\log\left(\frac{L(t,T_\delta^*)}{L(t,T_{2\delta}^*)}\right) + \frac{\delta L(t,T_\delta^*)}{1+\delta L(t,T_\delta^*)}\sigma^2(T_\delta^*)(T_\delta^* - t)}{\sqrt{v(t,T_{2\delta}^*,T_\delta^*)}} + \frac{1}{2}\sqrt{v(t,T_{2\delta}^*,T_\delta^*)} \tag{8.51}$$

$$d_2 = d_1 - \sqrt{v(t,T_{2\delta}^*,T_\delta^*)} \tag{8.52}$$

$$v(t,T_{2\delta}^*,T_\delta^*) = \sigma^2(T_\delta^*)(T_\delta^* - t) - 2\rho\sigma(T_\delta^*)\sigma(T_{2\delta}^*)\sqrt{(T_\delta^* - t)(T_{2\delta}^* - t)} + \sigma^2(T_{2\delta}^*)(T_{2\delta}^* - t) \tag{8.53}$$

$$\rho = \sqrt{\frac{T_{2\delta}^* - t}{T_\delta^* - t}} . \tag{8.54}$$

We now examine how well the approximations perform. Figure 5 plots the Monte Carlo estimates and the two approximations. Figure 5 shows that Approximation 1 performs very well. Compared to the Monte Carlo estimates, Approximation 1 underprices the resettable caplets by between 1% and 6%.

Approximation 2 does not perform so well and it overprices the resettable caplets by up to 5% and underprices by up to 24%.

Approximation 2 is a little disappointing but it is easy to implement. However, resettable caps are quoted by the average price of the constituent resettable caplets. Figure 6 plots the average resettable caplet prices as given by Approximation 2 and the Monte Carlo Simulations. Comparing the average prices, Approximation 2 underprices by the Monte Carlo prices up to 4% and overprices by up to 5%. Approximation 2 may perhaps be acceptable for providing indicative prices for resettable caps. Approximation 1 performs better but requires a proper calibration of the market-Libor model. We proceed to examine the calibration.

8.7 CALIBRATION ISSUES

The upper and lower bounds of Section 8.3 and the Approximation 1 of Section 8.6 all depend on the terms

$$\int_{t}^{t+k\delta} \underline{\lambda}(u,t+(k+1)\delta)\cdot\underline{\lambda}(u,t+k\delta)du\ ,\ k=1,\ \dots\ ,\ n, \tag{8.55}$$

where n is the number of resettable caplets making up the longest resettable cap we need to price. The numerical results of Section 8.5 were produced assuming we knew the forward Libor rate volatility factors, $\underline{\lambda}(u, T)$. In practice, we would have to calibrate the market model and extract the volatility factors, $\underline{\lambda}(u, T)$, from market data.

Forward-forward caplet volatilities only give the following constraints

$$\int_{0}^{k\delta} \left|\underline{\lambda}(u,k\delta)\right|^2 du = \sigma^2(k\delta)k\delta\ ,\ k=1\ ,\ \dots\ ,\ n+1 \tag{8.56}$$

on the forward Libor volatility factors: $\sigma(k\delta)$ is the forward-forward volatility of a standard caplet covering the period $[k\delta, (k+1)\delta]$ and n is the number of resettable caplets making up the longest resettable cap we need to price. The constraints

given by equation 8.56 permit considerable remaining flexibility for the choice of $\underline{\lambda}(u, T)$. Cap prices are not sensitive to the correlation structure of forward Libor rates and so they do not provide enough information to complete the terms in equation 8.55.

The calibration needs more market data. We may attempt to extract $\underline{\lambda}(u, T)$ by supplementing cap prices with historically estimated correlation of changes to forward Libor rates. For example, it may be desirable to assume the forward Libor volatility factors are time stationary so that $\underline{\lambda}(u, T) \equiv \underline{\lambda}(T-u)$. Then we can estimate an correlation matrix of forward Libor rate changes to provide the following additional constraints on the volatility factors

$$\frac{\underline{\lambda}(i\delta)\cdot\underline{\lambda}(j\delta)}{|\underline{\lambda}(i\delta)||\underline{\lambda}(j\delta)|} = \rho_{ij}\,,\ i, j = 1, \ldots, n, \tag{8.57}$$

where ρ_{ij} is the historically estimated correlation between proportional changes of the forward δ-Libor rates of maturities $i\delta$ and $j\delta$.

Calibrating the market-Libor model to caplet forward-forward volatilities and correlations is relatively easy. It is far easier than the calibration of conventional models where cap prices and correlations are highly non-linear functions of the model parameters and where it is difficult to understand how the model parameters affect the fit. Here, for the market-Libor model, it is very clear what the constraints on the volatility factors are.

Alternatively we may supplement the cap prices with market prices of options that are sensitive to the correlation structure of the forward Libor rates. Swaptions are often used because their market is liquid. However, calibrating to swaptions is more difficult because their values are highly non-linear with respect to the volatility factors, $\underline{\lambda}(u, T)$. The calibration would require a difficult non-linear optimisation. Swaptions cannot be priced analytically, but Brace, Gatarek and Musiela (1995) provide an approximate formula that performs well for their term

structures. This allows them to calibrate the market-Libor model simultaneously to caps and swaptions prices and a historically estimated correlation matrix.

The lack of an analytical pricing formula for swaptions point to a problem of the market-Libor model. The market-Libor model losses its analytical tractability when it is necessary to price options that depend on other non-Libor financial variables. The market-Libor models are only convenient for applications involving Libor rates. Other financial variables can be complex functions of the Libor rates that are intractable and difficult to simulate. It is probably worth mentioning here that Jamshidian (1996) has produced a market-swap model where forward swaps rates can be made lognormal for well chosen numéraires. In the market-swap model European swaptions are priced consistently with the market convention. Jamshidian (1996) shows that forward Libor rates and forward swap swaps cannot both be lognormal, that is, the market-Libor and market-swap models are inconsistent.

8.8 SUMMARY

The market Libor model offers some tractability for pricing Libor derivatives. We have derived an exact lower bound and an approximate upper bound for the prices of resettable caps and floors. We have also derived an exact relationship between resettable caplet and floorlet prices within the market-Libor model that depend only observables and forward-forward caplet volatilities that can be readily extracted from quoted cap volatilities. It would be interesting to test the relationship on quoted prices.

We derived two approximations for the prices of resettable caplets and floorlets. We examined the performance of the approximations using realistic Libor volatility factors. We found that the first approximation gives good approximations to the no-arbitrage prices. The second approximation does not perform so well but it is only a function of the current forward Libor term structure and caplet forward-

forward volatilities that are readily available to traders. The second approximation has the advantage that it be implemented easily in a spreadsheet to price approximately resettable caps and floors quickly off quotes for standard caps and floors.

We have provided a preliminary discussion on the calibrating issues related to the market-Libor model. We suggested that the market-Libor model would be easy to calibration to quoted cap volatilities and historically estimated forward Libor rate correlation. We argued that the market-Libor model losses some of its attraction when it is necessary to price or to calibrate it to non-Libor derivatives. More empirical work needs to be done to examine whether the market-Libor model is more suitable to practitioners than other interest rate term structure models.

8.9 APPENDIX

Lemma: Let $(X, Y) \sim N\left[\begin{pmatrix}\mu_x \\ \mu_y\end{pmatrix}, \begin{pmatrix}\sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2\end{pmatrix}\right]$, then

$$E\left[e^X - e^Y\right]^+ = \exp\left(\mu_x + \frac{\sigma_x^2}{2}\right)\phi\left[\frac{\mu_x - \mu_y + \sigma_x(\sigma_x - \rho\sigma_y)}{\sqrt{\sigma_x^2 - 2\rho\sigma_x\sigma_y + \sigma_y^2}}\right] - \exp\left(\mu_y + \frac{\sigma_y^2}{2}\right)\phi\left[\frac{\mu_x - \mu_y + \sigma_y(\sigma_y - \rho\sigma_x)}{\sqrt{\sigma_x^2 - 2\rho\sigma_x\sigma_y + \sigma_y^2}}\right]$$

where $\phi(.)$ is the normal distribution function.

8.10 REFERENCES

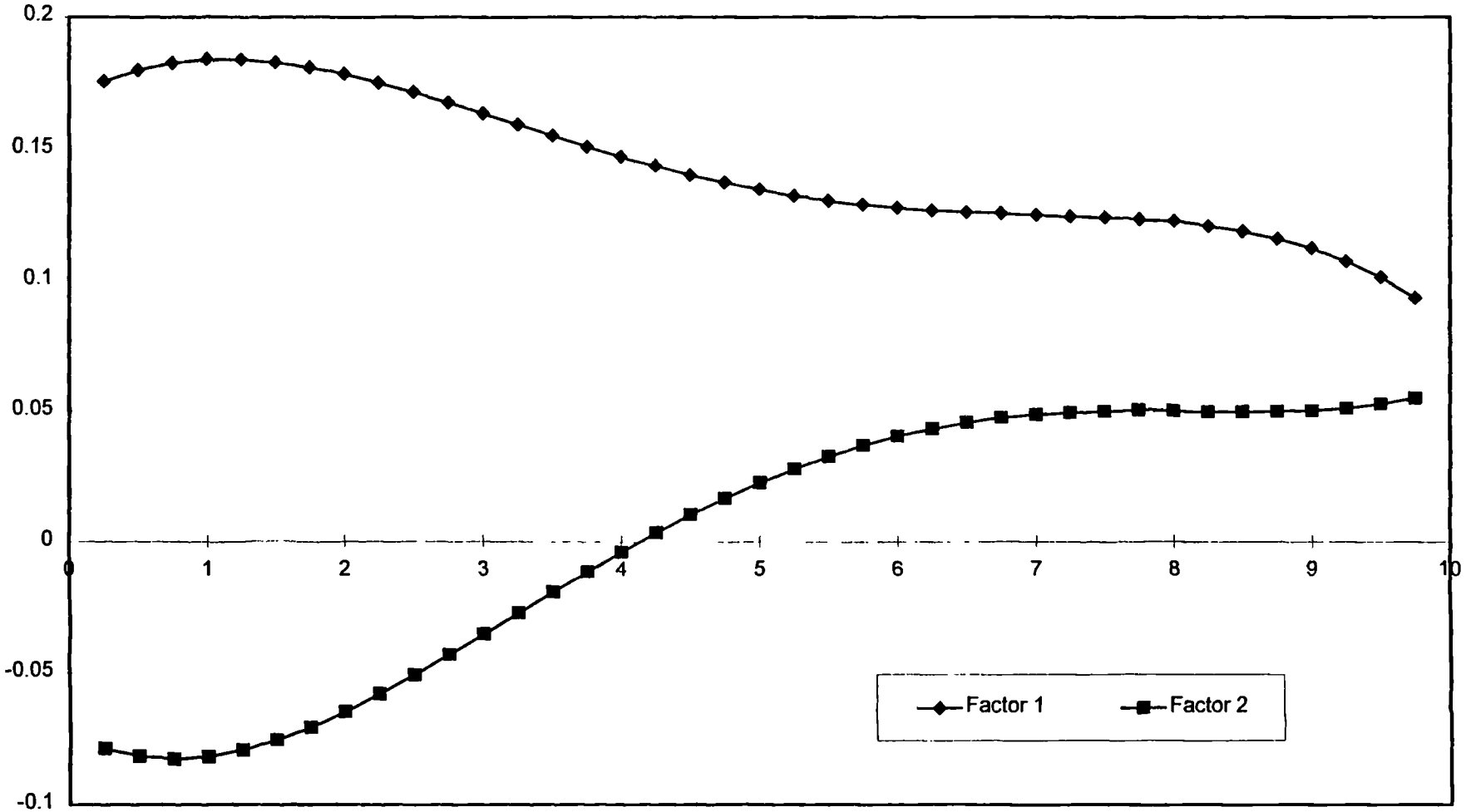
1. Brace A., Gatarek D. and Musiela M., "*The Market Model of Interest Rate Dynamics*", 1995, Working Paper, *USNW.*
2. Geman H., El Karoui N. and Rochet J-C, "*Changes of Numéraire, Changes of Probability Measure and Option Pricing*", 1995, Journal of Applied Probability.
3. Flesaker B. and Hughston L.P., "*Positive Interest: Foreign Exchange*", 1996, Vasicek and Beyond: Approaches to Building and Applying Interest Rate Models, (L.P. Hughston, ed., Risk Publications).

4. Harrison J.M. and Kreps D.M., 1979, "Martingales and Arbitrage in Multiperiod Securities Markets", *J. Econ. Theory* 20, 381-408.

5. Jamshidian F, "*Libor and Swap Market Models and Measures*", January 1996, Working Paper, Sakura Global Capital.

6. Musiela M and Rutkowski M, "*Continuous-time Term Structure Models*", September 1995, Working Paper, University of New South Wales.

Figure1: Proportional Forward Libor Volatility Factors



Forward Libor Maturity

Figure 2: Forward Libor and Volatility Term Structures

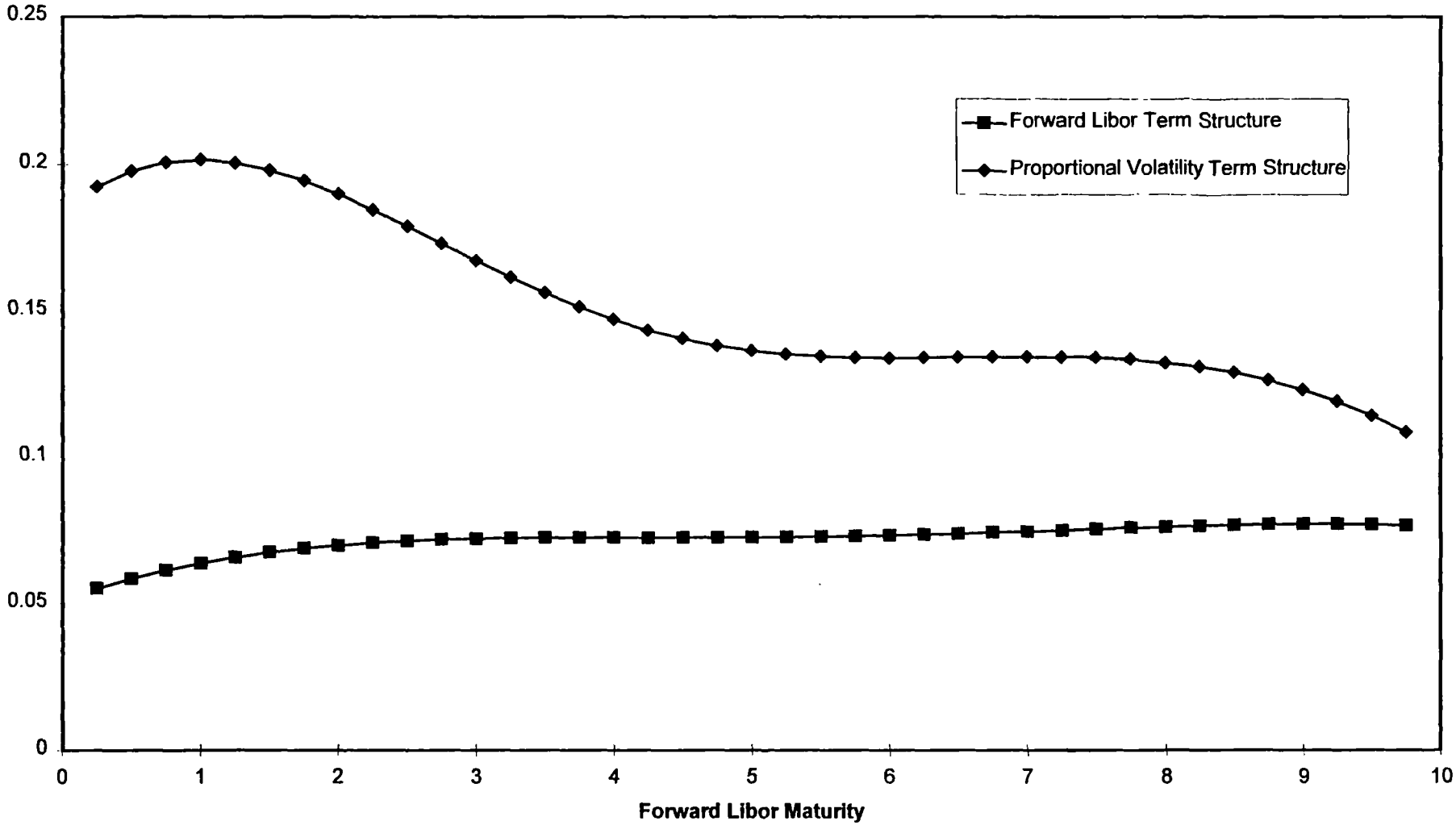


Figure 3: Upper and Lower Bounds to Resettable Caplet Value

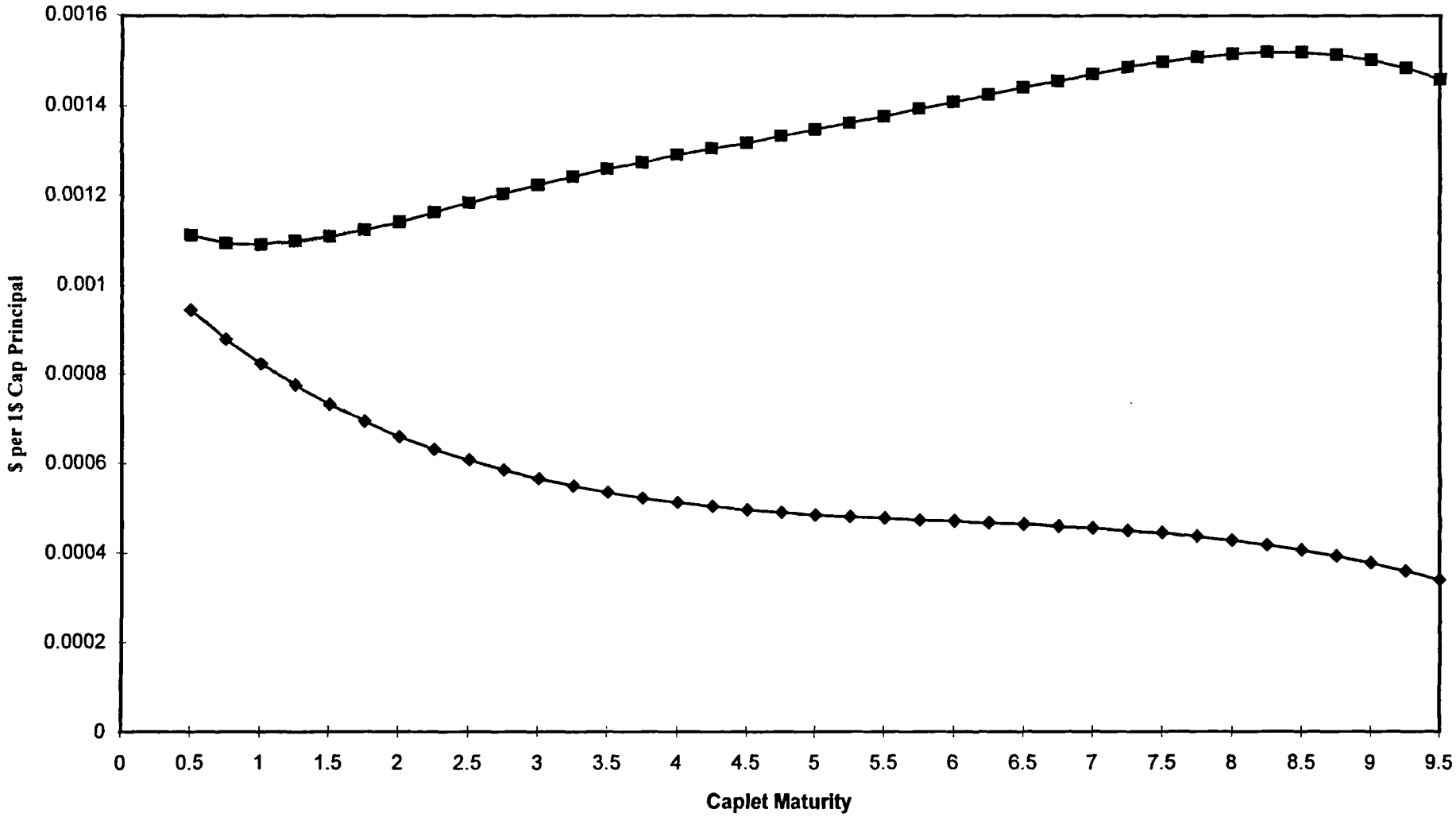


Figure 4: The Lower Bound and The No-Arbitrage Prices

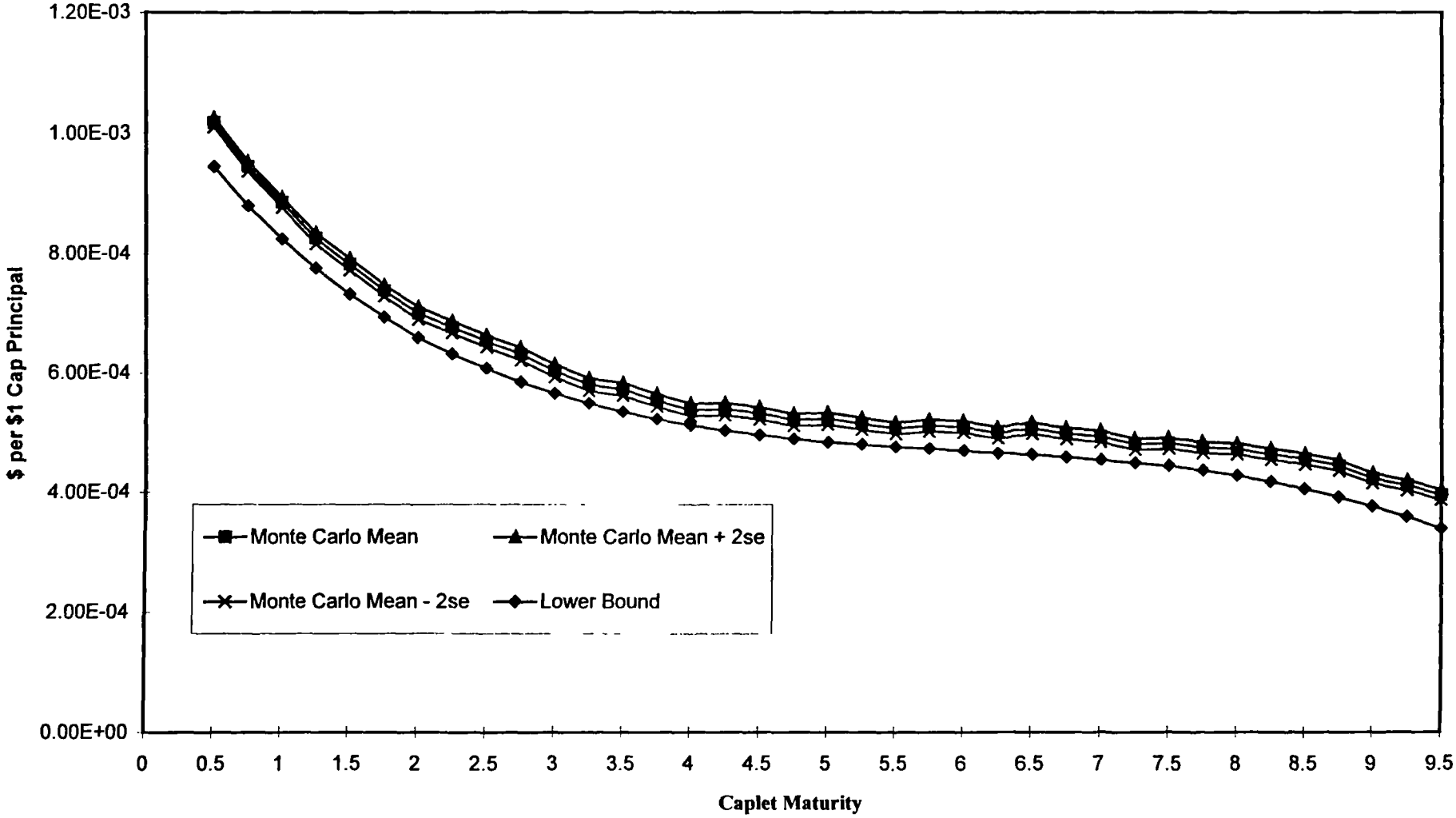


Figure 5: Approximations and No-Arbitrage Resettable Caplet Prices

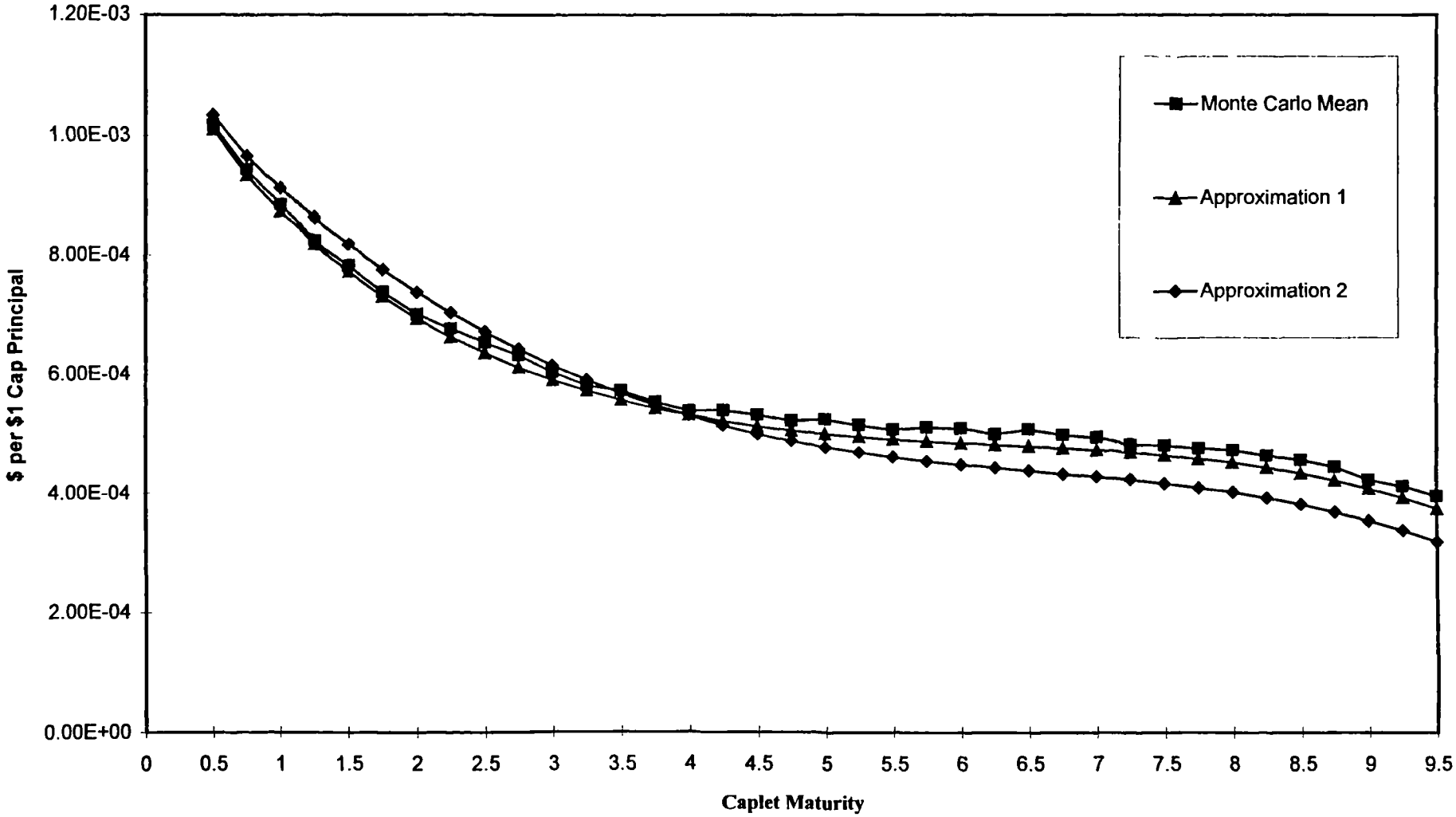
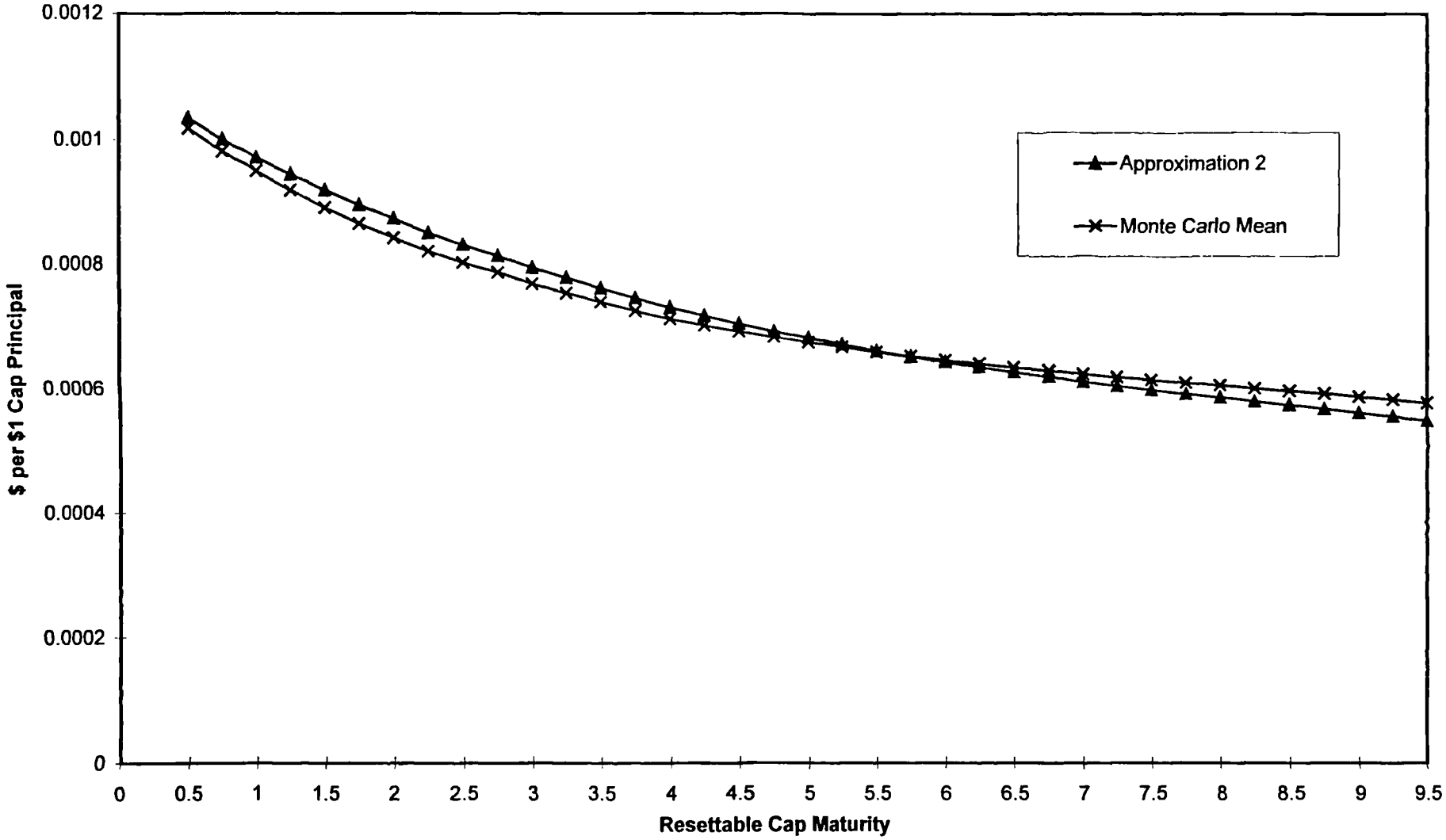


Figure 6: Average Resettable Caplet Prices



9. SUMMARY AND FURTHER RESEARCH

9.1 SUMMARY

We had set out to investigate interest rate term structure modelling and interest rate derivative pricing with particular emphasis on the calibration issues.

We argued in Chapter 2 that calibration issues are central to modelling considerations for if a model cannot be calibrated efficiently and quickly then it will have little practical value. There is a compromise between the ease with which a model can be calibrated and the accuracy that a model is able to capture empirical dynamics. We argued that the suitability of different models depend on their eventual use and in particular the type of derivatives the models will be used to price and hedge. After considering general modelling issues we provided a summary of the contingent claims pricing methodology and a review of the two major classes of interest rate term structure modelling models; Markovian short rate models and the Heath Jarrow and Morton models. Finally we provided an introduction to the numerical techniques that can be applied to valuation problems when analytical solutions are not available.

In Chapter 3 we provided a literature review of calibration work published in the literature. We found very few published papers on calibration issues. We found that even those papers that do calibrate a model do not discuss the suitability of their methods and do not test whether their fitted parameters are stable between re-calibrations.

In Chapter 4 we examined the calibration of short rate models. We reviewed well known construction techniques for Black, Derman and Toy (1990) and Hull and White (1993) trees. We showed how those techniques can be used to calibrate short rate trees to an interest rate term structure and a volatility term structure. We argued that the calibrated models will not be able to price even simple options

consistently with the market and that the models should be calibrated to options prices. We provided two examples by calibrating the how short rate trees to caplets and floorlets prices. The examples provided a simple setting for us to discuss many calibration issues. We proceeded to examine the properties of the calibrated short rate models and to discuss their misspecification. Finally we discussed the calibration of multifactor short rate models.

In Chapter 5 we examined a special class of multifactor short rate models; the Duffie and Kan Affine Yield Model where the state variables are zero coupon yields. We introduced the Duffie and Kan Affine Yield Model by discussing why multifactor short rate models are difficult to calibrate and how the Affine Yield Model seems to address each of the problems we raised. We examined the Duffie and Kan Affine Model and argued that practitioners would probably only be interested in the subclass that preclude negative interest rates, the Duffie and Kan Non-Negative Affine Yield Model. We provided an example to discuss how Affine Yield Models might be implemented and calibrated. We concluded from our example that Affine Yield Models, positive zero coupon yields or otherwise, are exceedingly difficult to calibrate. We then showed that the Duffie and Kan Non-Negative Affine Yield Models are equivalent to Generalised Cox, Ingersoll and Ross models that are relatively far easier to calibrate. This being so, we argued practitioners probably should calibrate Duffie and Kan Non-Negative Affine Yield Models using their Generalised Cox, Ingersoll and Ross equivalents. Thus Duffie and Kan Non-Negative Affine Yield Models do not make short rate models simpler to easier to calibrate.

In Chapter 6 we moved away from the short rate models to begin our analysis of Heath Jarrow and Morton type models. We reviewed how HJM forward rate trees can be used to price interest rate derivatives. We used a forward rate tree to price simple options and found that forward rate trees are far too slow to be

useful for calibration. We also considered pricing with Monte Carlo simulations. We discussed how the Monte Carlo simulations can be made more efficient and introduced the Martingale Variance Reduction technique of Carverhill and Pang (1995) to increase the accuracy of the Monte Carlo estimates. We found Monte Carlo simulations would also be too slow to be useful for calibrating HJM models. We demonstrated that we cannot calibrate HJM models quickly without restricting our attention to special subclasses characterised by their volatility structure. We finished Chapter 6 by examining two special subclasses; Gaussian HJM models and the Markovian HJM models of Cheyette (1992) and Ritchken and Sankarasubramanian (1995).

In Chapter 7 we introduced a new technique for calibrating Gaussian HJM model. We argued that we can use Kennedy (1994) as an intermediate step for calibrating Gaussian HJM. We began the chapter by reviewing two conventional approaches to the calibration and explaining why they are unsatisfactory. We examined in detail the Gaussian Random Field Model of Kennedy (1994). We showed how other Gaussian interest rate models are special cases of Kennedy (1994) and derived conditions that guarantees the continuity and smoothness of the instantaneous forward rate term structure. We showed how we can extract a Gaussian HJM approximation to Kennedy. We calibrated the Kennedy model can to caps and swaptions prices and extracted an Gaussian HJM approximation. We provide results that suggest the resulting Gaussian HJM model has better properties than those produced by conventional calibrations.

In Chapter 8 we examined the recently developed market-Libor model. After an introduction, we examined how the market-Libor model can be used to price resettable caps and floors to illustrate some of the advantages the market-Libor model offers for pricing Libor derivatives. We derived an exact relationship between the prices of resettable caplets and floorlets. We also derived an exact

lower bound, approximate upper bound and two approximations for the prices of resettable caps and floors. Using representative forward rate term structures volatility factors we demonstrated that the first approximation provide prices that are very close to the no-arbitrage prices of the caplets. The second approximation does not perform so accurately but has the advantage that it prices the resettable caps and floors directly of market variables. The second approximation can be implemented on a spreadsheet to provide indicative prices for resettable caps and floors. We finished the chapter by discussing how the market-Libor model might be calibrated.

9.2 FURTHER RESEARCH

We have argued in our thesis that for practical applications HJM type models are preferred to short rate models except for those derivatives that would be extremely difficult to price in the HJM approach. However, we have also seen that HJM models are difficult to calibrate except for the special subclasses that restrict the volatility factor we may use. This may pose severe problem for practitioners because it is highly probable that those instruments that we find difficult to price in the HJM approach are particular sensitive to the specification of the volatility factors. Further research would reveal the extent of this problem.

We have not examined models that allow interest rates to jump. The presence of jumps is supported by empirical data and the volatility skews and smiles that are quoted for interest rate derivative prices. However, models typically lose their analytic tractability when jumps are permitted. Recent jump-diffusion models typically adopt the no-arbitrage framework and can be separated along a similar line as we have in our thesis: The published papers generally extend existence models by allowing jumps in the short rate or the entire forward rate curve. From the discussions we have had already provided in our thesis, we can probably conclude the short rate models will be unsatisfactory if we have to allow a

second time dependent parameter to ensure a good fit and the whole curve models will be very difficult to implement and calibrate. So, we see again that there is compromise between tractability and realism of the model. Unless we can calibrate jump-diffusion interest rate models, practitioners are unlikely to find them suitable in application. Rather, practitioners will make do with easy to use models that the adjust to take into account their deficiencies. Much recent needs to be done to provide tractable jump-diffusion models.

We think that some of the most interesting and important issues to be addresses are misspecification problems. By misspecification, we mean the inability of the models to match empirical interest rate dynamics accurately. Since contingent claims pricing are based on the ability to form perfect dynamic hedges, misspecification problems are obviously very important. We have provided only a preliminary examination in Chapter 4. We saw that in some situations calibration and re-calibration can compensate for model misspecifications. It is important to investigate misspecification problems in greater depth and to see whether after allowing re-calibration the models can replicate the payoffs from a wide variety of interest rate derivatives.

10. GLOSSARY OF TERMS

Absolute Volatility	The volatility of the level of a random variable.
Forward Rate	The interest rate you can contract now to borrow or lend for a specified period, the duration, commencing at some specified time in the future, the maturity.
Implied Parameters	The parameters that have been extracted from market prices to allow the model to produce model prices that are close, according to some defined measure, to market prices.
Instantaneous Forward Rate	The forward rate that at maturity permits borrowing or lending for an instant.
Interest Rate Term Structure	The plot of interest rates, (zero coupon yields, forward rates, instantaneous forward rates, etc.), across their maturities. The type of interest rate will be clear from the context..
Interest Rate Volatility Term Structure	The plot of interest rate, (zero coupon yields, forward rates, instantaneous forward rates, etc.), volatilities across their maturities. The type of interest rate will be clear from the context..

Number of factors	Denotes the number of sources of uncertainty used to drive the interest rate dynamics. Thus single factor models of the interest rate term structure will have only one source of uncertainty that simultaneously affects all interest rates.
Numéraire	A numéraire is a price process that is almost surely positive.
Proportional Volatility	The volatility of the percentage change of a random variable.
Pure Discount Bond	A security that has no intermediate payments and pays unit cash at its maturity.
Short Rate	The Zero Coupon Yield on a pure discount bond that matures an instant later
Short Rate Model	A model in which the short rate process follows a Markovian process and in which the number of states variables is the same as the number of factors.
State Variables	Variables that characterise completely the state of the model so that all prices and all future distributions will be functions of the current levels of the state variables.

Volatility	Volatility, in diffusions, measures the rate uncertainty in random variables develop. Volatility is the square root of the variance rate.
Volatility Factor	The plot of the effect on interest rates, (zero coupon yields, forward rates, instantaneous forward rates, etc.), across their maturities of a unit change in a Brownian motion.
Zero Coupon Yield	The annualised continuously compounded return from holding a PDB until maturity. Zero coupon yields are referred to by the maturities of the pure discount bonds from which they are defined.

11. NOTATION

$d\underline{\widetilde{W}}_t$	Time t increment of $\underline{\widetilde{W}}_t$.
$d\underline{W}_t^T$	Time t increment of $\underline{W}_t^T$.
$\widetilde{E}_t[\cdot]$	Expectation conditional on the information available up to time t using the risk-neutral measure.
$E_t^T[\cdot]$	Expectation conditional on the information available up to time t using the probability measure that makes prices relative to the T maturity pure discount bond martingales.
$f(t, T)$ or $F(t, T)$	The time t instantaneous forward rate with maturity T.
$L(t, T)$	The time t forward δ-Libor rate with maturity T.
$L(T)$ or $L(T, T)$	The spot (time T) δ-Libor rate.
$P(t, T)$	The time t value of a pure discount bond that matures at time T paying unit value.
$\sigma_Y(\tau)$	The absolute volatility of the τ maturity zero coupon yield.
$\underline{\widetilde{W}}_t$	A Wiener process in the risk-neutral measure.

$\underline{W}_t^T$	A Wiener process in the T-measure induced by taking the T maturity PDB for the numéraire.
$Y(t, T)$	The zero coupon yield of a T maturity pure discount bond at time t.

12. BIBLIOGRAPHY

Adler R.J., 1981, *The Geometry of Random Fields*, New York: Wiley.

Amin K. and Morton A.J., "*Implied Volatility Functions in Arbitrage-Free Term Structure Models*", Journal of Financial Economics, 1994, Vol. 35, pp 141-180.

Black F. and Karasinski P., "*Bond and Option Pricing when Short Rates are Lognormal*", Financial Analysts Journal, July-Aug 1991, pp 52-59.

Black F., Derman E. and Toy W., "*A One-Factor Model of Interest Rates and It's Application to Treasury Bond Options*", Financial Analysts Journal, Jan-Feb 1990, pp 33-39.

Black F., March 1976, "*The Pricing of Commodity Contracts*", Journal of Financial Economics 3.

Brace A. and Musiela M., "*A Multifactor Gauss Markov Implementation of Heath, Jarrow and Morton*", Mathematical Finance, July 1994, Vol. 4(3), pp 259-283.

Brace A., Gatarek D. and Musiela M., "*The Market Model of Interest Rate Dynamics*", Working Paper, May 1995, University of New South Wales.

Broadie M. and Glasserman P., "*Pricing American-Style Securities Using Simulation*", Working Paper, 1994, Columbia University.

Brown R. and Schaefer S., "*Interest Rate Volatility and the Term Structure of Interest Rates*", Philosophical Transactions of the Royal Society: Physical Sciences and Engineering, 1993, Vol. 347: pp563-576.

Carverhill A.P. and Pang K., "*Efficient and Flexible Bond Option Valuation in the Heath, Jarrow and Morton Framework*", Journal of Fixed Income, September 1995, Vol. 5 No 2, pp. 70-77.

Carverhill A.P., "*Interest Rate Option Valuation Models: a Note about the Behaviour of the Volatility Structure*", Working Paper, 1992, Hong Kong University of Science and Technology.

Carverhill A.P., "*A Note on the Models of the Hull and White for Pricing Options on the Term Structure*", Journal of Fixed Income, September 1995.

Carverhill A.P., "*When is the Short Rate Markovian?*" Mathematical Finance, Vol. 4, 1994.

Chan K.C., Karolyi G.A., Longstaff F.A., and Sanders A.B., "*An Empirical Comparison of Alternative Models of the Short Term Interest Rate*", Journal of Finance, Vol. XLVII, No. 3, pp 1209-1227.

Chen L., "*Stochastic Mean and Stochastic Volatility: A Three-Factor Model of the Term Structure and its Application in Pricing in Interest Rate Derivatives*", Working Paper, Federal Reserve Board, October 1994.

Chen R.R. and Scott L., "*Pricing Interest Rate Options in a Two-Factor Cox-Ingersoll-Ross Model of the Term Structure.*", The Review of Financial Studies, Vol. 5 1992, pp. 613-636.

Cheyette O., "*Term Structure Dynamics and Mortgage Valuation*", Journal of Fixed Income, March 1992.

Clewlow L.J. and Carverhill A., "*On the Simulation of Contingent Claims*", Journal of Derivatives, Winter 1994, pp. 66-74.

Clewlow L.J., "*Finite Difference Techniques for One and Two Dimension Option Valuation Problems*", Nov 1992, FORC pre-print 90/10, University of Warwick.

Cox J.C., Ingersoll J.E. and Ross S.A., "*A Theory of the Term Structure of Interest Rates.*", Econometrica, March 1985, pp. 385-407.

Cox J.C., Ingersoll J.E. and Ross S.A., "*An Intertemporal General Equilibrium of Asset Prices*", Econometrica, Vol. 53, No.2, 1985.

Cox J.C., Ross S.A. and Rubinstein M., "*Option Pricing: A Simplified Approach*", Journal of Financial Economics, 1979, pp 229-263.

Duffie D. and Kan R., "*A Yield-Factor Model of Interest Rates.*", Mathematical Finance, 1996, Vol. 6, No. 4, pp 379-406.

El Karoui N., Geman H. and Lacoste V., "*On the Role of State Variables in Interest Rate Models*", Working Paper, May 1995, Université Paris VI, Laboratoire de Probabilités.

Flesaker B. and Hughston L.P., "*Positive Interest: Foreign Exchange*", 1996, Vasicek and Beyond: Approaches to Building and Applying Interest Rate Models, (L.P. Hughston, ed., Risk Publications).

Fong H.G. and Vasicek O., "*Interest Rate Volatility as a Stochastic Factor*", Working Paper, Gifford Fong Associates, 1991.

Geman H., El Karoui N. and Rochet J-C., 1995, "*Changes of Numéraire, Changes of Probability Measure and Option Pricing*", Journal of Applied Probability, Vol. 32, pp. 443-458.

Gibbons M.R. and Ramaswamy K, "*A Test of the Cox, Ingersoll and ross Model of the Term Structure*", Review of Financial Studies, Vol. 6(3), 1993, pp 619-58.

Harrison J.M. and Kreps D.M., 1979, "Martingales and Arbitrage in Multiperiod Securities Markets", *J. Econ. Theory* 20, 381-408.

Harrison J.M. and Pliska S.R., 1981, "Martingales and Stochastic Integrals in the Theory of Continuous Trading", *Stoch. Proc. Appl.* Vol. 11, pp. 215-260.

Heath D., Jarrow R. and Morton A., "*Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation*", Econometrica, Vol. 60, 1992, pp. 77-105.

Heath D., Jarrow R. and Morton A., "*Contingent Claim Valuation with a Random Evolution of Interest Rates*", Journal of Futures Markets, Vol. 10, 1991.

Ho T.S.Y. and Lee S.B., "*Term Structure Movements and Pricing Interest Rate Contingent Claims*", Journal of Finance, 1986, Vol. XLI, No. 5, pp 1011-1029.

Hull J. and White A., "*One Factor Interest Rate Models and the Valuation of Interest Rate Derivative Securities*", Journal of Financial and Quantitative Analysis, Vol. 28 (2), June 1993, pp 235-254.

Hull J. and White A., "*Pricing Interest Rate Derivative Securities*", The Review of Financial Studies, Vol. 3 (4), 1990.

Hull J. and White A., "*'A Note on the Models of Hull and White for Pricing Options on the Term Structure': Response*", The Journal of Fixed Income, September 1995, pp 97-102.

Hull J. and White A., "*Efficient Procedures for Valuing European and American Path-Dependent Options*", The Journal of Derivatives, Fall 1993, pp 21-31.

Hull J. and White A., "*Numerical Procedures for Implementing term Structure Models II: Two-Factor Models*", The Journal of Derivatives, Winter 1994, pp 37-48.

Hull J. and White A., "*Using Hull-White Interest Rate Trees*", The Journal of Derivatives, Spring 1996, pp 26-36.

Jamshidian F, "*Libor and Swap Market Models and Measures*", January 1996, Working Paper, Sakura Global Capital.

Jamshidian F., "*Forward Induction and Construction of Yield Curve Diffusion Models*", Merrill Lynch working paper, 1991.

Jamshidian, F., "*An Exact Bond Option Formula*", Journal of Finance, March 1989, Vol. XLIV, No.1 , pp 205-209.

Joy C., Boyle P. and Tan K.S., "*Quasi Monte Carlo Methods in Numerical Finance*", Working paper, February 1995, University of Waterloo.

Kennedy D.P., 1995, "*Characterizing and Filtering Gaussian Models of the Term Structure of Interest Rates*", Working Paper, University of Cambridge.

Kennedy D.P., July 1994, "*The Term Structure of Interest Rates as a Gaussian Random Field*", Mathematical Finance, Vol 4 No 3, pp. 247-258.

Langetieg T.C., "*A Multivariate Model of the Term Structure.*", Journal of Finance, March 1980, pp. 71-97.

Longstaff F.A. and Schwartz E.D., "*Interest-Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model.*", Journal of Finance, September 1992, pp. 1259-1282.

Musiela M and Rutkowski M, "*Continuous-time Term Structure Models*", September 1995, Working Paper, University of New South Wales.

Niederreiter H., "*Numerical Integration III*", ISNM vol. 85, Bräss H. and Hämmerlin G., eds. (Basel: Birkhaüser), pp. 157-171.

Paskov S.H., "*Computing High Dimensional Integrals with Applications to Finance*", Technical Report CUCS-023-94, October 1994, Columbia University.

Press, Teukolsky, Vetterlin and Flannery, "*Numerical Recipes in C*", Second Edition, Cambridge University Press.

Rebonato R. and Cooper I., 1996, "*The Limitations of Simple Two-Factor Interest Rate Models*", Journal of Financial Engineering, Vol 5 No 1.

Ritchken P. and Sankarasubramanian L., "*Volatility Structures of Forward Rates and the Dynamics of the Term Structure*", Mathematical Finance, Vol. 5 No. 1.

Smith G.D., "*Numerical Solution of Partial Differential Equations*", Oxford Mathematical Books, Oxford University Press, 1975.

Vasicek O., 1977, "*An Equilibrium Characterization of the Term Structure*", J. Financial Econ., 5, pp. 177-188.