Essays on Stochastic Volatility and Random-Field Models in Finance

HELEN TSOULOUVI

Thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Economics

LONDON SCHOOL OF ECONOMICS

2005

Abstract

In this thesis we develop random-field models for the implied volatility of equity options and the term structure of interest rates. Following a brief introduction to the topics of this thesis in chapter 1, chapter 2 models the Black-Scholes implied volatility of plain-vanilla European stock options as a random field with three parameters: current time, the maturity date and the exercise price of the corresponding option. In this model all plain-vanilla European options are needed to complete the market. Illiquid and exotic derivatives can be priced as a function of the stock price and the implied volatility surface.

In chapter 3 we develop a random-field model for forward interest rates with stochastic volatility. It is assumed that the forward rate volatility function can be decomposed into a deterministic function of the time to maturity and a maturity-independent stochastic process driven by a standard Brownian motion. The separability of the forward-rate volatility function allows closed-form solutions to be obtained for the prices of a number of interest rate derivatives: bond options, interest rate caplets, and interest rate spread options. Forward LIBOR and swap rates are modelled in a similar way, and closed-form solutions are derived for the prices of LIBOR caplets and swaptions.

In chapter 4 we estimate three random-field models of the term structure of interest rates: one model with deterministic forward-rate volatility, and two with stochastic forward-rate volatility. The models are estimated using seven years of daily UK and US forward rate data, spanning times to maturity between zero and 120 months. The parameters of each model are obtained by maximizing the likelihood function. We develop an importance sampling technique that substantially reduces the variance of the Monte Carlo estimator of the likelihood function in the case of stochastic volatility.

Acknowledgements I wish to thank Antonio Mele, Angelos Dassios, and Vassilis Hajivassiliou for helpful comments and suggestions, and my parents for their support and encouragement.

Contents

List of Tables

Chapter 1

Introduction

Random field models were introduced in finance by Kennedy (1994, 1997) for the purpose of modelling forward interest rates and bond prices. This class of models provides a natural extension of the popular Heath, Jarrow, and Morton (HJM) (1992) term structure model to infinite factors. A random field is a multiparameter stochastic process. An n-parameter random field is a stochastic process with random innovations that vary with respect to each of the n parameters. By analogy with the physical system described by Brownian motion – a particle moving randomly in space – a two-parameter random field can be thought of as the random motion of a string, and a three-parameter random field as the random motion of a membrane.[1] Random field term structure models offer greater flexibility relative to the HJM framework, while at the same time the number of model parameters can be kept low, by assuming, for example, that the correlation structure of the random field shocks is a smooth function of the time to maturity.

Random field models may also be employed to model the Black-Scholes implied volatility of plain-vanilla European stock options. There exists abundant evidence, which is discussed in the following chapters, that the volatility of both stock and bond prices is stochastic, i.e. it cannot be written as a function of the asset price and time only. In the case of stock options, implied volatility can be used to obtain information about the unobservable stock price volatility process. In similarity to the HJM model, which fits the initial forward rate curve, one can build a model that fits the initial implied volatility surface for a given stock. Using the dynamics of implied volatility, illiquid and exotic derivatives can be priced, taking the prices of liquid European options as given.

[1] In this physical analogy the first parameter of the random field represents time, other parameters correspond to the position of a point on the string or the membrane, and the value of the random field may indicate velocity, for example.

In this thesis we develop stochastic volatility models that incorporate random field innovations for the purpose of pricing stock and interest rate derivatives. The case of stock options is considered in the next chapter, where we model the implied volatility of plain-vanilla European stock options as a random field with three parameters: current time, the maturity date and the exercise price of the corresponding option. The no-arbitrage condition on asset prices implies that the drift of implied volatility is endogenously determined. Using this drift restriction, we show that implied volatility converges to a measure of the instantaneous volatility of the underlying asset price, as the time to maturity tends to zero, and the exercise price tends to the current stock price. In this model markets are incomplete due to stochastic volatility in asset prices, in contrast to the standard Black-Scholes model. We show how the prices of illiquid and exotic derivatives can be obtained as a function of the stock price and implied volatility. The random-field implied volatility model allows greater flexibility in fitting observed volatility smiles than previously studied stochastic volatility models, none of which has been found to fit observed patterns in implied volatility adequately.[2]

In chapter 3 we develop a random field model for forward interest rates with stochastic volatility. We assume that the forward rate volatility function can be decomposed into a deterministic function of the time to maturity and a maturity-independent stochastic process driven by a standard Brownian motion. The latter can be correlated with the random field shocks driving forward rates. The separability of the forward-rate volatility function allows us to obtain closed-form solutions for the prices of a number of standard interest rate derivatives: bond options, interest rate caplets, and interest rate spread options. In practice, the most common interest rate contingent claims are LIBOR[3] and swap rate derivatives. As a result, LIBOR and swap rate models are more popular with market practitioners than models for continuously compounded forward rates. We develop random field models for both forward LIBOR and swap rates with a separable volatility structure, and derive closed-form solutions for the prices of LIBOR caplets and swaptions.

Chapter 4 provides an empirical study of random field term structure models. Three models are considered. In the first model forward rate volatility is a function of the spot rate and the time to maturity. In the second model forward rate volatility is a function of the time to maturity and a stochastic process driven by a standard Brownian motion. In the third model volatility is driven by a random field. In both the second and third

[2]See for example Das and Sundaram (1999). We discuss the empirical evidence in chapter 2.
[3]London Inter-Bank Offer Rate.

models it is assumed that volatility innovations are uncorrelated with the random field process driving forward rates. We estimate the models using seven years of daily UK and US forward rate data, spanning times to maturity between zero and 120 months, at increments of one month. We use all maturities available in our dataset, in order to approximate as closely as possible the continuous nature of random fields.

Previous empirical studies of random field term structure models (Longstaff, Santa-Clara, and Schwartz, 2001a and 2001b) reduce the model to a discrete version with a small number of factors – only four in the above papers. This makes the model indistinguishable from finite-factor models, as shown by Kerkhof and Pelsser (2002). In fact very few empirical studies of high-dimensional models exist, due to the technical difficulties present in such situations, especially in the case of models with latent variables. While the number of empirical studies of stochastic volatility models has increased rapidly in recent years, most of the econometric methods developed so far are either inefficient or computationally unfeasible when applied to high-dimensional latent-variable models.

We estimate the three models by maximizing the likelihood function of the observed data. In the case of stochastic volatility we perform simulations to integrate volatility out of the likelihood function, a technique known as Monte Carlo maximum likelihood. We develop a method of importance sampling that substantially reduces the variance of the Monte Carlo estimator of the likelihood function, so that the number of required simulations is low enough to make the estimation process feasible.

Chapter 2

A random-field model for implied volatility

2.1 Introduction

The celebrated Black-Scholes formula (Black and Scholes, 1973) is widely used by market participants as a benchmark for pricing options. The shortcomings of the model and its discrepancies with option price data are well documented. The model assumes that returns on the underlying asset are normally distributed and that the volatility, or instantaneous variance, of the asset price is constant. However, empirical studies of stock returns find that unconditional distributions have tails which are thicker than those of the normal distribution, and that volatility is time-varying and exhibits persistence of shocks. Some of the earliest evidence is given by Mandelbrot (1963) and Fama (1965).[1] Volatility is also inversely related to the stock price level – the 'leverage effect' – (Beckers, 1980), and exhibits mean reversion in the long run (Merville and Pieptea, 1989).

The Black-Scholes model can be used to solve for the implied volatility of an option by inverting the pricing formula. In practice, plots of implied volatility against the exercise price usually form a convex curve (the volatility smile, smirk or skew depending on whether it is symmetric or not) or, less often, a concave curve (the volatility frown). Furthermore, implied volatility can be increasing or decreasing in the time to maturity. These observations provide further evidence against the validity of the Black-Scholes model's assumption of constant volatility.[2] Numerous models which dispense with

[1] See also Blattberg and Gonedes (1974), Engle and Mustafa (1992), and Bollerslev, Chou, and Kroner (1992).

[2] Apart from nonconstant volatility, other possible explanations for the observed deviations of option price data from the Black-Scholes model are stochastic interest rates, transaction costs, liquidity problems, and feedback effects from dynamic hedging.

this assumption have been proposed. Such models can be arranged in two groups: deterministic volatility (DV) models, where volatility is a function of observable variables, and stochastic volatility (SV) models, where the stochastic process driving volatility introduces one or more additional risk factors in the economy.

The earliest DV model (Merton, 1973) extended the Black-Scholes model so that volatility is a deterministic function of time. Cox and Ross (1976) consider a model where volatility is proportional to the level of the stock price raised to a power – the constant elasticity of variance (CEV) diffusion model. Geske (1979), Rubinstein (1983), and Bensoussan, Crouhy, and Galai (1994) present DV models where option prices are a function of the value of the firm.

Alternative DV models have been developed in discrete time through the implied tree approach and GARCH (generalized autoregressive conditional heteroskedasticity) time series modelling. GARCH option pricing models (Engle and Mustafa, 1992; Duan, 1995; Kallsen and Taqqu, 1998) assume that current volatility is a weighted average of past squared asset returns and past volatility, with greater weight being assigned to more recent values. Thus GARCH models can capture volatility clustering – the tendency of high (or low) volatility values to persist in time. GARCH option pricing models do not have closed-form solutions however. As a result, simulation methods are usually employed to compute option prices and hedge ratios.[3]

Implied tree DV models (Derman and Kani, 1994; Rubinstein, 1994; Dupire, 1994 and 1997) assume that asset price movements form a binomial (or trinomial) tree. Under this hypothesis the implied risk-neutral distribution of the asset price at maturity is obtained from the observed prices of European options. Volatility is assumed to be a deterministic function of the asset price and time, but no assumptions are made about its functional form, which may be quite complex and unique for every different ending return distribution. Models of this class fit initial option prices but are not very successful in predicting future option prices (Dumas, Fleming, and Whaley, 1998).

An important difference between DV and SV models concerns market completeness. In DV models options are redundant securities while SV models introduce additional risk factors in the economy. As a result, options are not redundant securities in SV models, unless one of the following two assumptions is made:

(1) volatility is a traded asset, *or*

(2) aggregate consumption is uncorrelated with volatility, so that the risk premium asso-

[3]A related type of DV models are exponentially weighted moments models, proposed by Hobson and Rogers (1998). This class of models assumes that instantaneous volatility is a function of exponentially weighted moments of the historic logarithms of stock prices.

ciated with volatility under the physical measure is zero.

If either (1) or (2) holds, options can be valued using the unique risk-neutral equivalent martingale measure. Otherwise, markets (excluding options) are incomplete, and there exist an infinite number of equivalent martingale measures, with each one corresponding to a particular specification for the volatility risk premium.

A popular type of SV models are diffusion models, which are of the following form in general:[4]

$$\begin{aligned} dS(t) &= \mu S(t)\,dt + \sigma(t)\,S(t)\,dW_S(t) && (2.1)\\ d\sigma(t) &= a[t,\sigma(t)]\,dt + b[t,\sigma(t)]\,dW_\sigma(t) && (2.2) \end{aligned}$$

where $S(t)$ denotes the stock price and $\sigma(t)$ its instantaneous volatility. $W_S(t)$ and $W_\sigma(t)$ are standard Brownian motions with correlation coefficient ρ:

$$cor\left[dW_S(t), dW_\sigma(t)\right] = \rho$$

Table (2.1) shows different specifications for the functions a and b in the volatility diffusion equation employed in a number of SV models. All of the models in the table use equation (2.1) as the diffusion equation for stock prices, except Johnson and Shanno (1987), who modify this equation to:

$$dS(t) = \mu S(t)\,dt + \sigma(t)\,S^{\gamma}(t)\,dW_S(t)$$

Model	$a[t,\sigma(t)]$	$b[t,\sigma(t)]$
Wiggins (1987)	$f[\sigma(t)]$	$\beta\sigma(t)$
Scott (1987), Stein and Stein (1991)	$\delta[\vartheta - \sigma(t)]$	β
Johnson and Shanno (1987)	$\eta\sigma(t)$	$\beta\sigma^{\zeta}(t)$

Table 2.1: Volatility diffusion equation parameters

Hull and White (1987), and Heston (1993) model the dynamics of the square of volatility, $v(t) = \sigma^2(t)$, as follows:

$$\begin{aligned} dv(t) &= \eta v(t)\,dt + \beta v(t)\,dW_v(t) && \text{(Hull and White, 1987)}\\ dv(t) &= \varkappa[\theta - v(t)]\,dt + \beta\sqrt{v(t)}dW_v(t) && \text{(Heston, 1993)} \end{aligned}$$

[4]Nelson (1990) and Duan (1996, 1997) show that SV diffusion models can be obtained as the continuous-time limit of GARCH models.

Wiggins (1987) makes specific assumptions about preferences and model parameters, such as the asset's beta coefficient in the capital asset pricing model, and the partial correlation coefficient between the return on the market and the individual stock's volatility. He uses finite differences numerical methods to solve the partial differential equation for option prices. Johnson and Shanno (1987) assume that there exists an asset whose price process is perfectly correlated with volatility. They price stock options using Monte Carlo simulations.

Hull and White (1987), Scott (1987), and Stein and Stein (1991) assume that $\rho = 0$, and that the volatility risk premium is either zero or equal to a constant. Hull and White (1987) obtain a series approximation for the price of a European call option.[5] Scott (1987) prices options using Monte Carlo simulations. Stein and Stein (1991) derive a closed-form solution for the stock price distribution. A drawback of the Scott (1987), and Stein and Stein (1991) models is that the mean-reverting Ornstein-Uhlenbeck process used to model $\sigma(t)$ can become negative.

Heston (1993) obtains closed-form solutions for European option prices in terms of the inverse Fourier transform of the characteristic function of the stock price logarithm. His model allows ρ to take any value. Thus it can account for the observed negative correlation between asset returns and volatility. The volatility risk premium is assumed to be proportional to $\sigma^2(t)$. Schobel and Zhu (1999) use Heston's method to price options in the model by Stein and Stein (1991) with non-zero correlation between returns and volatility. Bakshi, Cao, and Chen (1997), and Scott (1997) also derive option prices using Heston's method for SV models which include random jumps in returns, and stochastic interest rates.

A number of general equilibrium SV models have also been proposed (Bailey and Stulz, 1989; Amin and Ng, 1993; Bakshi and Chen, 1997; Garcia, Luger, and Renault, 2002). Equilibrium models, however, have the disadvantage that option prices depend on the model's assumptions about preferences. Our survey of SV models is by no means exhaustive. For example, other researchers have considered SV models that can replicate particular aspects of volatility behaviour, such as random jumps resulting from the discontinuous arrival of information (see for example Naik, 1993), and long memory (Comte and Renault, 1998).

Empirical research suggests that SV models explain option prices better than DV models. Dumas, Fleming, and Whaley (1998) compare the performance of a number of DV models against an ad hoc (and internally inconsistent, but often used in practice)

[5] They also present a numerical solution for the case where $\rho \neq 0$.

option pricing procedure where Black-Scholes implied volatilities are smoothed across exercise prices and maturities, and are used as an input in the Black-Scholes model. They find that DV models perform no better in hedging or predicting option prices than this ad hoc procedure. Their results imply that although the Black-Scholes model is inconsistent with the data, DV models are unable to improve on option pricing accuracy. Bates (1997) compares three types of models: CEV (DV) models, SV models with negative correlation between volatility and returns, and jump-diffusion models with negative-mean jumps in returns. He finds that CEV models are less consistent with options market data than the other two types of models.

Bakshi, Cao, and Chen (2000), and Buraschi and Jackwerth (2001) provide evidence that options are not redundant securities, particularly after the 1987 crash, and that additional risk factors need to be added to one-dimensional diffusion models to explain the data. Lamoureux and Lastrapes (1993) provide evidence against SV models with zero volatility risk premium. Their data suggest that the volatility risk premium is time-varying and a decreasing function of the volatility level. Pan (2002) and Benzoni (2002) also find that the volatility risk premium is statistically significant. A study of the performance of delta-hedged S&P 500 index option portfolios by Bakshi and Kapadia (2003) suggests that the market volatility risk premium is negative.

Using S&P 500 index option price data, Nandi (1996) finds that Heston's (1993) model performs better than the Black-Scholes model in terms of pricing and hedging. In a subsequent paper Nandi (1998) finds evidence of nonzero correlation between returns and volatility using Heston's (1993) model. Bakshi, Cao, and Chen (1997), and Buraschi and Jackwerth (2001) find that one-dimensional diffusion models of asset returns are not improved significantly by the inclusion of stochastic interest rates, relative to the inclusion of stochastic volatility. Jones (2003) observes that as volatility increases, its own volatility increases at an even faster rate.[6]

Recently, market models of implied volatility have been proposed by Schonbucher (1999) and Ledoit, Santa-Clara, and Yan (2002), in similarity to market models of the term structure of interest rates, introduced by Brace, Gatarek, and Musiela (1997) and Jamshidian (1997)[7]. In this type of models the Black-Scholes implied volatility (IV) is assumed to follow a continuous-time process with specific dynamics. Ledoit *et al.* (2002) model the dynamics of IV, denoted by $V(t,s,X)$, as follows:

$$dV(t,s,X) = \mu_V(t,s,X)\,dt + \sigma_{V_1}(t,s,X)\,dW_1(t) + \sigma_{V_2}(t,s,X)\,dW_2(t)$$

[6]Garcia, Ghysels and Renault (2004) provide a survey of recent empirical work on stochastic volatility.
[7]Market models of the term structure of interest rates are discussed in section 3.3.

where t denotes current time, s the time to maturity of the option, and X the *moneyness* of the option, defined as the stock price divided by the exercise price. W_1 and W_2 are independent Brownian motions, with W_1 being the Brownian motion driving the stock price process in equation (2.1). The functions $\mu_V(t,s,X)$, $\sigma_{V_1}(t,s,X)$, and $\sigma_{V_2}(t,s,X)$ can depend on IV and the stock price in general. Using the discretized dynamics of the stock price and a Taylor approximation to the standard normal distribution function, Ledoit *et al.* (2002) show that as the time to maturity goes to zero, the IV of at-the-money options converges to the instantaneous volatility of the stock price process. Using this result, the instantaneous volatility in equation (2.1) can be replaced by $V(t,0,1)$. Another important result in this framework is that the IV drift is endogenously determined through the partial differential equation that a contingent claim must satisfy and the Black-Scholes partial differential equation, through which IV is defined.

Schonbucher (1999) presents a similar IV model. The dynamics of the stock price and IV under the risk-neutral measure are assumed to be as follows:

$$\begin{aligned} dS(t) &= rS(t)\,dt + \sigma(t)\,S(t)\,dW_0(t) \\ dV(t,T,K) &= u(t,T,K)\,dt + \gamma(t,T,K)\,dW_0(t) + \sum_{n=1}^{N} v_n(t,T,K)\,dW_n(t) \end{aligned}$$

where T denotes the maturity date and K the exercise price of the option, W_n $n = 0,..,N$ are Brownian motions, and W_0 is orthogonal to W_n for $n = 1,..,N$. When $N = 1$ the model reduces to the Ledoit *et al.* (2002) model. Brace *et al.* (2002) show that these IV models encompass all Markovian stochastic instantaneous volatility models. Brace *et al.* (2001) apply this approach to the modelling of the IV of forward LIBOR caplets.[8, 9]

It would be natural to extend the Schonbucher (1999) and Ledoit *et al.* (2002) models so that IV is driven by a random field with three parameters – current time, the maturity date of the corresponding option, and its exercise price – as suggested by Ledoit *et al.* (2002). When IV is driven by a random field, the number of sources of stochastic shocks in options markets is infinite.[10] In this case IV and, hence, option prices are infinite-dimensional stochastic processes, i.e. they cannot be spanned by a finite number of independent stochastic processes. It follows that all options are needed to complete the market in this framework.

[8]The implied volatility of interest rate caplets is calculated using the Black (1976) model (see chapter 3).

[9]Another type of models that fit initial IV are implied tree models, which were discussed in the context of DV models. Derman and Kani (1998), and Britten-Jones and Neuberger (2000) extend the implied tree approach to include stochastic volatility. Skiadopoulos (2001) provides a survey of this class of models.

[10]We exclude degenerate cases where the random field collapses to a finite number of Brownian motions.

This result, however, is inconsistent with simple option pricing models, such as those by Hull and White (1987) and Heston (1993). In these two models, for example, option prices are spanned by two stochastic processes, and market completeness is achieved using two securities, such as the stock and an option, or two options (Bajeux-Besnainou and Rochet, 1996; Romano and Touzi, 1997). We therefore need to examine whether there exists any specification for the dynamics of the stock price and its instantaneous volatility that is compatible with IV being modelled as an (infinite-dimensional) random field. If such a specification does not exist, then the assumption that IV is a random field necessarily implies that the no-arbitrage condition on option prices does not hold, possibly due to market imperfections, such as transaction costs and liquidity problems.

The structure of the rest of this chapter is as follows. In section 2.2 the stochastic IV model is developed and we address the above consistency issues. Under the assumptions of the model, the entire continuum of plain-vanilla European options for all maturity dates and exercise prices is required to complete the market. The drift of the IV process is endogenously determined, as a result of the no-arbitrage condition on option prices and the Black-Scholes differential equation, through which IV is defined. Using the IV drift restriction equation we show that IV converges to a measure of the instantaneous volatility of the stock price, as the time to maturity tends to zero and the exercise price tends to the current stock price. In section 2.3 we price volatility derivatives using Heston's (1993) inverse Fourier transform approach. In section 2.4 we show how the prices of illiquid and exotic derivatives can be obtained using the dynamics of IV. Finally, section 2.5 concludes.

2.2 Stochastic implied volatility model

Uncertainty is modelled as a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{P}$ represents the physical measure, and $\mathbb{F}$ is assumed to satisfy the *usual conditions*.[11] There exists a non-dividend-paying stock with price process $S(t)$. Options can be written on this stock for all exercise prices K and maturity dates T, such that $0 < K < \infty$ and $t < T < \infty$. There also exists a riskless asset, or bank account, with price process $B(t)$. We assume that the instantaneous nominal rate of interest r is a positive constant, and the initial value of the bank account, $B(0)$, is equal to one. Thus, the value of the bank account at time t is $B(t) = e^{rt}$ and its dynamics are:

$$dB(t) = rB(t)\,dt \tag{2.3}$$

[11] $\mathcal{F}_0$ contains all the null sets of $\mathbb{P}$, and $\mathbb{F}$ is right-continuous.

The Black-Scholes formula for the price of a European call option $C(t)$ with exercise price K and maturity date T at time t, written on the stock $S(t)$, is given by

$$C(t) = S(t)\,N(d_1) - e^{-r(T-t)} K N(d_2) \tag{2.4}$$

where $N(x)$ is the standard normal distribution function, and d_1, d_2 are equal to:

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{1}{2}v^2\right)(T-t)}{v\sqrt{T-t}} \qquad d_2 = d_1 - v\sqrt{T-t} \tag{2.5}$$

where v is the instantaneous volatility of the stock price, assumed to be constant, in the Black-Scholes model.

The implied volatility of an option with price C is defined as the value of v for which the Black-Scholes formula generates an option price equal to C. It must be emphasized that the use of Black-Scholes implied volatilities does not imply that the Black-Scholes model is valid. It is merely a one-to-one relationship between a measure of volatility and the market price of an option with a given maturity date and exercise price. For this reason options are usually quoted in terms of their IV instead of their market price. We model IV because it is a directly observable quantity, in contrast to the instantaneous volatility of the stock price process. Furthermore, this approach allows us to fit the initial market prices of plain-vanilla European options. In practice v is a stochastic process which depends on t, T, and K. Hence it can be represented as a random field $v(t,T,K)$. This random field can be visualized as a surface or membrane evolving stochastically through time.

We define the following function

$$G_{TK}(S,v,t) = S(t)\,N(d_1) - e^{-r(T-t)} K N(d_2) \tag{2.6}$$

which gives the price of a European call option with maturity date T and exercise price K, as a function of S, v, and t.

The Black-Scholes formula is the solution to the following partial differential equation:

$$\frac{\partial G_{TK}}{\partial t} + \frac{1}{2}v^2 S^2 \frac{\partial^2 G_{TK}}{\partial S^2} + rS\frac{\partial G_{TK}}{\partial S} - rG_{TK} = 0 \tag{2.7}$$

subject to the boundary condition $G_{TK}(T) = \max[S(T) - K, 0]$.

The partial derivatives of G_{TK}, denoted by Greek letters by convention, are as follows:

$$\begin{aligned}
\Delta &= \frac{\partial G_{TK}}{\partial S} = N(d_1) \qquad\qquad \Gamma = \frac{\partial^2 G_{TK}}{\partial S^2} = \frac{n(d_1)}{Sv\sqrt{T-t}} \\
\Theta &= \frac{\partial G_{TK}}{\partial t} = -\frac{Sn(d_1)\,v}{2\sqrt{T-t}} - rKe^{-r(T-t)} N(d_2)
\end{aligned}$$

$$\nu = \frac{\partial G_{TK}}{\partial v} = S\sqrt{T-t}n(d_1)$$

$$\frac{\partial^2 G_{TK}}{\partial v^2} = S\sqrt{T-t}n(d_1)\frac{d_1 d_2}{v} \qquad \frac{\partial^2 G_{TK}}{\partial S \partial v} = -\frac{d_2 n(d_1)}{v}$$

where $n(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ is the standard normal density function.

The IV of a put option is equal to the IV of a call option with the same exercise price and expiration date. This follows from the relationship below, known as the put-call parity:

$$C(t) - P(t) = S(t) - Ke^{-r(T-t)}$$

where $P(t)$ denotes the price of a European put option with the same maturity date and exercise price as $C(t)$. This relationship holds irrespective of the model that describes the dynamics of the stock price and volatility, under the assumption that there do not exist any arbitrage opportunities.

We specify the dynamics of $v(t,T,K)$ as follows:

$$d_t v(t,T,K) = \alpha(t,T,K)\,dt + \beta(t,T,K)\,d_t Z(t,T,K) \tag{2.8}$$

where $Z(t,T,K)$ is a random field with three parameters, and the subscript t in the differential operator indicates that the change in the random field is obtained with respect to changes in parameter t, holding T and K constant.

ASSUMPTION I *The random field $Z(t,T,K)$ satisfies the following conditions:*

(i) $Z(t,T,K)$ *is continuous in* t, T, K.

(ii) $E[d_t Z(t,T,K)] = 0$.

(iii) $var[d_t Z(t,T,K)] = dt$.

(iv) $cor[d_t Z(t,T_1,K_1), d_t Z(t,T_2,K_2)] = c(T_1,K_1,T_2,K_2)$.

(v) $d_t Z(t_1,T_1,K_1)$ *and* $d_t Z(t_2,T_2,K_2)$ *are independent* $\forall\ t_1 \neq t_2$.

Under the above conditions the correlation structure of the random field $Z(t,T,K)$ does not depend on t. From (2.6) we obtain that the dynamics of $C(t)$ satisfy the following equation, using Ito's lemma:

$$\begin{aligned} dC &= \frac{\partial G_{TK}}{\partial t}dt + \frac{\partial G_{TK}}{\partial S}dS + \frac{1}{2}\frac{\partial^2 G_{TK}}{\partial S^2}d\langle S\rangle + \frac{\partial G_{TK}}{\partial v_{TK}}d_t v_{TK} \\ &+ \frac{1}{2}\frac{\partial^2 G_{TK}}{\partial v_{TK}^2}d_t\langle v_{TK}\rangle + \frac{\partial^2 G_{TK}}{\partial S \partial v_{TK}}d_t\langle v_{TK}, S\rangle \end{aligned} \tag{2.9}$$

where angles $\langle\ \rangle$ denote the quadratic variation of a process or the quadratic covariation of two processes, and $v_{TK}(t) = v(t,T,K)$.

Let $\sigma(t)$ denote the instantaneous volatility of the stock price. We assume that $\sigma(t)$ is stochastic, but we leave it unspecified for the moment. Option prices can be written as a function of the stock price and its instantaneous volatility, as a result of the no-arbitrage condition that they must satisfy. Defining this function as $C(t) = H_{TK}(S,\sigma,t)$, the dynamics of $C(t)$ are also given by the following equation, from Ito's lemma:

$$dC = \frac{\partial H_{TK}}{\partial t}dt+\frac{\partial H_{TK}}{\partial S}dS+\frac{1}{2}\frac{\partial^2 H_{TK}}{\partial S^2}d\langle S\rangle+\frac{\partial H_{TK}}{\partial \sigma}d\sigma+\frac{1}{2}\frac{\partial^2 H_{TK}}{\partial \sigma^2}d\langle\sigma\rangle+\frac{\partial^2 H_{TK}}{\partial S\partial\sigma}d\langle\sigma,S\rangle \tag{2.10}$$

Let $\vartheta(t)$ denote the risk premium associated with $\sigma(t)$, and let $\mu(t)$ denote the drift of $S(t)$. The no-arbitrage condition on option prices is given by the following partial differential equation (Garman, 1976):

$$E(dC) = \left(rC + \vartheta\frac{\partial H_{TK}}{\partial\sigma} + S(\mu - r)\frac{\partial H_{TK}}{\partial S}\right)dt \tag{2.11}$$

where E is the expectation operator.

Since we have two alternative characterizations of option prices, we need to find out whether the above model is viable, i.e. whether there exist stock price and volatility processes that can generate option prices and IV which can be described as a random field. The answer is not obvious, because $S(t)$ and $\sigma(t)$ are not functions of T or K, and, as a result, any stochastic shocks appearing on the right hand side of equation (2.10) will be functions of t only, while equations (2.8) and (2.9) imply that option prices change randomly with respect to T and K. A way to make the two equations comparable arises if the random field can be written as the weighted sum of an infinite number of independent Brownian motions with weights that vary with respect to T and K.

Definition 2.1 *An infinite-dimensional Brownian motion is a sequence* $W(t) = \{W_i(t), i = 0,1,2,\ldots\}$ *of independent standard Brownian motions.*

ASSUMPTION II *The random field* $Z(t,T,K)$ *can be written in terms of an infinite-dimensional Brownian motion as follows:*

$$Z(t,T,K) = \sum_{i=0}^{\infty}\zeta_{TKi}W_i(t) \tag{2.12}$$

where ζ_{TKi} *is a function of* T, K *and* i.[12]

The dynamics of IV can thus be written as:

$$d_t v(t,T,K) = \alpha(t,T,K)\,dt + \beta(t,T,K)\sum_{i=0}^{\infty}\zeta_{TKi}dW_i(t) \tag{2.13}$$

[12]The model can be easily extended to include time-varying ζ_{TKi}.

We assume the following dynamics for the stock price and an associated sequence of volatilities $\sigma_i(t)$:

$$\begin{aligned} dS(t) &= \mu(t) S(t) dt + S(t) \sum_{i=1}^{\infty}\sum_{j=0}^{\infty} \kappa_{ij}\sigma_i(t) dW_j(t) &\quad (2.14)\\ d\sigma_i(t) &= \gamma_i(t) dt + \delta_i(t) dW_i(t) \qquad i = 1, 2, \ldots &\quad (2.15)\end{aligned}$$

where $\mu(t), \gamma_i(t)$ and $\delta_i(t)$ can be functions of $S(t)$, $\sigma_i(t)$, and t in general, and the κ_{ij} are constants.

ASSUMPTION III *$S(t)$, $\sigma_i(t)$, and $v(t,T,K)$ are strictly positive $\forall\ t,\ i,\ t \leq T < \infty$, $0 < K < \infty$. Furthermore, the parameters in the dynamics of $S(t)$, $\sigma_i(t)$, and $v(t,T,K)$ satisfy appropriate Lipschitz and linear growth conditions so that there exist unique and bounded strong solutions to the stochastic differential equations (2.13)–(2.15).*[13]

ASSUMPTION IV *(i) $\exists$ at least one i such that $\kappa_{i0} \neq 0$.*
(ii) $\delta_i(t) \neq 0\ \ \forall\ i,\ t$.
(iii) $\beta(t,T,K) \neq 0\ \ \forall\ t,\ t < T < \infty,\ 0 < K < \infty$.

Let $\alpha_{TK}(t) = \alpha(t,T,K)$ and $\beta_{TK}(t) = \beta(t,T,K)$. Using equations (2.13)–(2.15) we expand equations (2.9) and (2.10):

$$\begin{aligned} dC &= \left[\frac{\partial G_{TK}}{\partial t} + \frac{\partial G_{TK}}{\partial S}\mu S + \frac{\partial G_{TK}}{\partial v_{TK}}\alpha_{TK} + \frac{1}{2}\frac{\partial^2 G_{TK}}{\partial S^2}S^2\sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\sigma_i\kappa_{ij}\right)^2\right.\\ &\quad \left. +\frac{1}{2}\frac{\partial^2 G_{TK}}{\partial v_{TK}^2}\beta_{TK}^2\sum_{i=0}^{\infty}\zeta_{TKi}^2 + \frac{\partial^2 G_{TK}}{\partial S\partial v_{TK}}S\beta_{TK}\sum_{j=0}^{\infty}\zeta_{TKj}\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right]dt\\ &\quad +\sum_{j=0}^{\infty}\left(\frac{\partial G_{TK}}{\partial S}S\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i + \frac{\partial G_{TK}}{\partial v_{TK}}\beta_{TK}\zeta_{TKj}\right)dW_j \qquad (2.16)\\ dC &= \left[\frac{\partial H_{TK}}{\partial t} + \frac{\partial H_{TK}}{\partial S}\mu S + \sum_{i=1}^{\infty}\frac{\partial H_{TK}}{\partial \sigma_i}\gamma_i + \frac{1}{2}\frac{\partial^2 H_{TK}}{\partial S^2}S^2\sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\sigma_i\kappa_{ij}\right)^2\right.\\ &\quad \left. +\frac{1}{2}\sum_{i=1}^{\infty}\frac{\partial^2 H_{TK}}{\partial \sigma_i^2}\delta_i^2 + S\sum_{i=1}^{\infty}\frac{\partial^2 H_{TK}}{\partial S\partial\sigma_i}\delta_i\sum_{j=1}^{\infty}\kappa_{ji}\sigma_j\right]dt\\ &\quad +\frac{\partial H_{TK}}{\partial S}S\sum_{i=1}^{\infty}\kappa_{i0}\sigma_i dW_0 + \sum_{j=1}^{\infty}\left(\frac{\partial H_{TK}}{\partial S}S\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i + \frac{\partial H_{TK}}{\partial \sigma_j}\delta_j\right)dW_j \quad (2.17)\end{aligned}$$

Let $\mathbf{A}$ denote the infinite-dimensional matrix with ijth element $\zeta_{T_jK_ji}$, where $t < T_j < \infty$ and $0 < K_j < \infty$. When the columns of $\mathbf{A}$ are linearly independent, the

[13]These conditions, which depend on the functional forms of $\mu(t), \gamma_i(t), \delta_i(t), \alpha(t,T,K)$, $\beta(t,T,K)$, and ζ_{TKi}, can be found in Da Prato and Zabczyk (1992).

random field $Z(t,T,K)$ cannot be spanned by a finite number of independent stochastic processes. Let $\mathbf{C}(t)$ and $\boldsymbol{v}(t)$ denote two infinite-dimensional vectors consisting of option prices and implied volatilities, respectively, for all maturity dates $t < T < \infty$ and exercise prices $0 < K < \infty$, and let $\boldsymbol{\sigma}(t)$ and $\mathbf{W}(t)$ denote the vectors of all $\sigma_i(t)$ and Brownian motions respectively. The two mappings

$$\begin{aligned} \mathcal{G} : (S, \boldsymbol{v}) &\rightarrow (S, \mathbf{C}) \\ \mathcal{H} : (S, \boldsymbol{\sigma}) &\rightarrow (S, \mathbf{C}) \end{aligned}$$

must be consistent with each other. Under assumptions (III) and (IV) the mapping $\mathbf{W} \rightarrow (S, \boldsymbol{\sigma})$ is invertible $\forall\ t$. If the Jacobian of $\mathcal{H}$ is nonzero $\forall\ t$, then by the inverse function theorem $\mathcal{H}$ is also invertible, and hence all options (and the stock) are needed to complete the market.

Proposition 2.2 *Under assumptions (I)–(IV), if the Jacobian of $\mathcal{H}$ is nonzero $\forall\ t$, then the columns of $\mathbf{A}$ are linearly independent and the mapping $\mathbf{W} \rightarrow (S, \boldsymbol{v})$ is invertible.*

PROOF. The Jacobian of $\mathcal{H}$ is the determinant of the following matrix:

$$\mathbf{J} = \begin{pmatrix} 1 & \frac{\partial H_{T_1K_1}}{\partial S} & \frac{\partial H_{T_2K_2}}{\partial S} & \cdots & \frac{\partial H_{T_iK_i}}{\partial S} & \cdots \\ \mathbf{0} & & \mathbf{N} & & & \end{pmatrix}$$

where $\mathbf{0}$ is an infinite-dimensional column vector of zeros, and $\mathbf{N}$ is the matrix of the first-order partial derivatives of H with respect to all $\sigma_i(t)$, i.e. its ijth element is equal to $\frac{\partial H_{T_jK_j}}{\partial \sigma_i}$. Nonsingularity of $\mathbf{J}$ is equivalent to nonsingularity of $\mathbf{N}$. Equations (2.16) and (2.17) imply that

$$\begin{aligned} \frac{\partial H_{TK}}{\partial S} S \sum_{i=1}^{\infty} \kappa_{i0}\sigma_i &= \frac{\partial G_{TK}}{\partial S} S \sum_{i=1}^{\infty} \kappa_{i0}\sigma_i + \frac{\partial G_{TK}}{\partial v_{TK}} \beta_{TK}\zeta_{TK0} \\ \frac{\partial H_{TK}}{\partial S} S \sum_{i=1}^{\infty} \kappa_{ij}\sigma_i + \frac{\partial H_{TK}}{\partial \sigma_j}\delta_j &= \frac{\partial G_{TK}}{\partial S} S \sum_{i=1}^{\infty} \kappa_{ij}\sigma_i + \frac{\partial G_{TK}}{\partial v_{TK}} \beta_{TK}\zeta_{TKj} \quad j = 1, 2, \ldots \end{aligned}$$

or in matrix form:

$$\mathbf{GBA}' = \left(S \sum_{i=1}^{\infty} \kappa_{i0}\sigma_i \, \mathbf{y} \quad \mathbf{N}'\boldsymbol{\Delta} + S\mathbf{V} \right) \tag{2.18}$$

where $\mathbf{A}'$ denotes the transpose of $\mathbf{A}$; $\boldsymbol{\Delta}$, $\mathbf{B}$, and $\mathbf{G}$ are diagonal matrices with main diagonals consisting of all δ_i, $\beta_{T_iK_i}$, and $\frac{\partial G_{T_iK_i}}{\partial v_{T_iK_i}}$ respectively; $\mathbf{y}$ is an infinite-dimensional column vector with ith element equal to $\frac{\partial H_{T_iK_i}}{\partial S} - \frac{\partial G_{T_iK_i}}{\partial S}$, and $\mathbf{V}$ is an infinite-dimensional matrix with ijth element equal to

$$\left(\frac{\partial H_{T_iK_i}}{\partial S} - \frac{\partial G_{T_iK_i}}{\partial S} \right) \sum_{m=1}^{\infty} \kappa_{mj}\sigma_m.$$

$\mathbf{G}$ is nonsingular since $\frac{\partial G_{TK}}{\partial v_{TK}}$ is strictly positive for $t < T < \infty$ and $0 < K < \infty$ under assumption (III). $\mathbf{\Delta}$ and $\mathbf{B}$ are also nonsingular by assumption (IV). Using the following operations on the matrix on the right hand side of equation (2.18):

(a) dividing the first column by $\sum_{i=1}^{\infty} \kappa_{i0}\sigma_i$, and

(b) multiplying the first column by $\sum_{m=1}^{\infty} \kappa_{mj}\sigma_m$ and subtracting it from column $j+1$ for $j = 1, 2, \dots,$

we obtain the matrix $(S\mathbf{y} \;\; \mathbf{N}'\mathbf{\Delta})$. Hence, if $\mathbf{N}$ is nonsingular then the columns of $\mathbf{A}$ are linearly independent. $\mathcal{G}$ is always invertible since its Jacobian is equal to:

$$\begin{vmatrix} 1 & \frac{\partial G_{T_1K_1}}{\partial S} \quad \frac{\partial G_{T_2K_2}}{\partial S} \quad \cdots \\ \mathbf{0} & \mathbf{G} \end{vmatrix}$$

The following mapping is therefore feasible and invertible:

$$\mathbf{W} \to (S, \boldsymbol{\sigma}) \to (S, \mathbf{C}) \to (S, \boldsymbol{v}) \quad \square$$

Henceforth we shall assume the following:

ASSUMPTION V *The columns of* $\mathbf{A}$ *are linearly independent and the mapping* $\mathbf{W} \to (S, \boldsymbol{v})$ *is invertible.*

Under the above assumption, $\mathcal{H}$ is also invertible, all options are needed to complete the market, and we can price any derivative using $(S, \boldsymbol{v})$. Let $\mathbf{K}$ denote the infinite-dimensional matrix with ijth element equal to κ_{ij}. If the rank of $\mathbf{K}$ is finite, then the dynamics of the stock price can be written in terms of a finite number of stochastic processes that are linear combinations of the volatilities $\sigma_i(t)$ and the Brownian motions $W_i(t)$ and, as a result, the number of options required to complete the market is finite. Hence, for assumption (V) to hold, the rank of $\mathbf{K}$ must be infinite.

ASSUMPTION VI *There exists a stochastic discount factor* $M(t)$ *with the following dynamics:*

$$\frac{dM(t)}{M(t)} = -r dt - \sum_{i=0}^{\infty} \lambda_i(t)\, dW_i(t) \tag{2.19}$$

where $\lambda_i(t)$ *denotes the risk premium associated with Brownian motion* $W_i(t)$*.*

This risk premium can be a function of the stock price and $\sigma_i(t)$, in general. Under the assumption that there exist no arbitrage opportunities, the product of the discount factor with the price of a traded asset must be a martingale. The drift of the discount factor is $-r$, so that $M(t)B(t)$ is a martingale. By Ito's lemma, the dynamics of the product of the discount factor with the stock price are:

$$\frac{d(MS)}{MS} = \left(\mu - r - \sum_{j=0}^{\infty}\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\lambda_j\right)dt + \sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i - \lambda_j\right)dW_j$$

Hence, the drift of the stock price process must satisfy the following condition:

$$\mu = r + \sum_{j=0}^{\infty}\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\lambda_j \tag{2.20}$$

In the same way we obtain a restriction on the drift of option prices. Using Ito's lemma and equation (2.16), the dynamics of $M(t)\,C(t)$ are (we suppress the subscripts T and K):

$$\begin{aligned}\frac{d(MC)}{M} &= \left[\frac{\partial G}{\partial t} + \frac{\partial G}{\partial S}\mu S + \frac{1}{2}\frac{\partial^2 G}{\partial S^2}S^2\sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right)^2 + \frac{\partial G}{\partial v}\alpha\right.\\ &\quad + \frac{\partial^2 G}{\partial S\partial v}S\beta\sum_{j=0}^{\infty}\sum_{i=1}^{\infty}\kappa_{ij}\zeta_j\sigma_i + \frac{1}{2}\frac{\partial^2 G}{\partial v^2}\beta^2\sum_{i=0}^{\infty}\zeta_i^2 - rC\\ &\quad \left. - \sum_{j=0}^{\infty}\lambda_j\left(\frac{\partial G}{\partial v}\beta\zeta_j + \frac{\partial G}{\partial S}S\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right)\right]dt\\ &\quad + \sum_{j=0}^{\infty}\left[\left(\frac{\partial G}{\partial v}\beta\zeta_j + \frac{\partial G}{\partial S}S\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right) - \lambda_j C\right]dW_j\end{aligned}$$

For the drift in the above equation to be zero, the following must hold:

$$\begin{aligned}rC + \sum_{j=0}^{\infty}\lambda_j\left(\frac{\partial G}{\partial v}\beta\zeta_j + \frac{\partial G}{\partial S}S\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right) &= \frac{\partial G}{\partial t} + \frac{\partial G}{\partial S}\mu S + \frac{1}{2}\frac{\partial^2 G}{\partial S^2}S^2\sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right)^2\\ &\quad + \frac{\partial G}{\partial v}\alpha + \frac{1}{2}\frac{\partial^2 G}{\partial v^2}\beta^2\sum_{j=0}^{\infty}\zeta_j^2 + \frac{\partial^2 G}{\partial S\partial v}S\beta\sum_{j=0}^{\infty}\sum_{i=1}^{\infty}\kappa_{ij}\zeta_j\sigma_i\end{aligned}$$

The above equation is equivalent to the Garman (1976) equation (2.11), when option prices are expressed in terms of implied volatilities. Combining this equation with the Black-Scholes partial differential equation (2.7) we obtain:

$$\begin{aligned}\sum_{j=0}^{\infty}\lambda_j\left(\frac{\partial G}{\partial v}\beta\zeta_j + \frac{\partial G}{\partial S}S\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right) &= \frac{\partial G}{\partial S}S(\mu - r) + \frac{1}{2}\frac{\partial^2 G}{\partial S^2}S^2\left[\sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right)^2 - v^2\right]\\ &\quad + \frac{\partial G}{\partial v}\alpha + \frac{1}{2}\frac{\partial^2 G}{\partial v^2}\beta^2\sum_{j=0}^{\infty}\zeta_j^2 + \frac{\partial^2 G}{\partial S\partial v}S\beta\sum_{j=0}^{\infty}\sum_{i=1}^{\infty}\kappa_{ij}\zeta_j\sigma_i\end{aligned}$$

Substituting the Black-Scholes partial derivatives into the above equation, we obtain the following condition for the IV drift:

$$\begin{aligned}\alpha\sqrt{T-t} &= \frac{d_1}{2}\left(T - t - \frac{d_1\sqrt{T-t}}{v}\right)\beta^2\sum_{j=0}^{\infty}\zeta_j^2 - \sqrt{T-t}\,\beta\sum_{j=0}^{\infty}\zeta_j\left(\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i - \lambda_j\right)\\ &\quad + \frac{d_1}{v}\beta\sum_{j=0}^{\infty}\sum_{i=1}^{\infty}\kappa_{ij}\zeta_j\sigma_i - \frac{\sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right)^2 - v^2}{2v\sqrt{T-t}}\end{aligned} \tag{2.21}$$

Substituting for d_1 from equation (2.5), multiplying throughout by $2v\sqrt{T-t}$ and rearranging, we obtain:[14]

$$\begin{aligned}
v^2 &= \left[\ln\left(\frac{S}{K}\right)+\left(r+\frac{1}{2}v^2\right)(T-t)\right]\times \\
&\quad \left\{\frac{1}{v^2}\left[\ln\left(\frac{S}{K}\right)+\left(r+\frac{1}{2}v^2\right)(T-t)\right]-(T-t)\right\}\beta^2\sum_{j=0}^{\infty}\zeta_j^2 \\
&\quad -\frac{2}{v}\left[\ln\left(\frac{S}{K}\right)+\left(r+\frac{1}{2}v^2\right)(T-t)\right]\beta\sum_{j=0}^{\infty}\sum_{i=1}^{\infty}\kappa_{ij}\zeta_j\sigma_i \\
&\quad +2v(T-t)\beta\sum_{j=0}^{\infty}\zeta_j\left(\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i-\lambda_j\right)+2v\alpha(T-t)+\sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right)^2
\end{aligned}$$

Taking limits as $T\to t$ and $K\to S(t)$,

$$\lim_{\substack{T\to t\\ K\to S(t)}} v^2 \to \sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right)^2 \tag{2.22}$$

Thus we obtain the Ledoit *et al.* (2002) result for the case of an infinite number of volatility processes $\sigma_i(t)$.[15] The dynamics of $v[t,t,S(t)]$ can be obtained using the Ito-Venttsel formula (see appendix A):

$$\begin{aligned}
dv[t,t,S(t)] &= \left[\mu S\frac{\partial v}{\partial K}+\frac{1}{2}\frac{\partial^2 v}{\partial K^2}S^2\sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right)^2+S\sum_{j=0}^{\infty}\frac{\partial}{\partial K}\left(\beta_{tS}\zeta_{tSj}\right)\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right. \\
&\quad \left.+\alpha_{tS}+\frac{\partial v}{\partial T}\right]dt+\sum_{j=0}^{\infty}\left(\beta_{tS}\zeta_{tSj}+\frac{\partial v}{\partial K}S\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right)dW_j \qquad (2.23)
\end{aligned}$$

where the partial derivatives are calculated at $T=t$ and $K=S(t)$. In the models of Schonbucher (1999) and Ledoit *et al.* (2002) $v[t,t,S(t)]$ is equal to $\sigma(t)$, and hence the IV of at-the-money near-the-expiration European options can be used for pricing other derivatives. In our model $v[t,t,S(t)]$ is not useful for pricing derivatives, but it still provides a measure of the total instantaneous volatility of the stock price process, since $d\langle S(t)\rangle = S^2(t)\,v^2[t,t,S(t)]\,dt$.

2.3 Pricing volatility derivatives

A number of options exchanges have introduced volatility indices which are based on the implied volatility of options. For example, the Chicago Board Options Exchange

[14]Note that we assume that S, v, and σ_i are strictly positive and bounded stochastic processes $\forall\ t, i, T$, and K.

[15]Ledoit *et al.* (2002) prove that IV converges to the instantaneous volatility of the stock price using discrete-time approximations. Our proof does not rely on approximations, but it requires the assumption that $v[t,t,S(t)]$ is strictly positive.

(CBOE) introduced a volatility index, named VIX, in 1993. Originally VIX was constructed in such a way that its value represented the implied volatility of a hypothetical at-the-money S&P 100 index option with 30 days to maturity. Its calculation has recently been changed so that it includes out-of-the-money options and the underlying index is the S&P 500 index. In March 2004 trading in VIX futures began on the CBOE Futures Exchange, and there are plans to introduce VIX options in the near future.

Derivatives on volatility allow investors to hedge against and place bets on volatility. In the absence of volatility derivatives, it is possible, in principle, to trade volatility through a portfolio which is highly sensitive to volatility, such as a straddle, or through a dynamic trading strategy. In practice, however, such methods do not replicate volatility perfectly, while they may also present problems associated with discrete adjustments, transaction costs, and frequent monitoring needs.

Grunbichler and Longstaff (1996) present a model for valuing futures and options on volatility. They specify a mean-reverting diffusion process for the volatility V of returns on a stock index. V can alternatively represent IV. Although V is not a traded asset, it is assumed that it is possible to construct a non-self-financing portfolio which exactly replicates V. The volatility risk premium is assumed to be proportional to V. The price of a European call option on V is obtained in terms of non-central chi-squared distribution functions and a normal approximation is used to evaluate it. Detemple and Osakwe (2000) extend this model to cover more specifications for the volatility process, and value American as well as European options.

In this chapter we assume that there exists only one risky asset in the economy, the stock (excluding options). The model can be easily extended to include multiple risky assets driven by Brownian motions which are correlated with each other. We define the following stochastic process:

$$\frac{dm(t)}{m(t)} = \sum_{i=0}^{\infty} \lambda_i(t)\, dW_i(t) \tag{2.24}$$

Given the stochastic discount factor defined by equation (2.19), the change of measure from the physical to the risk-neutral is obtained from the following equation for each Brownian motion $W_i(t)$ (Ikeda and Watanabe, 1989, p. 192; Duffie, 1996):

$$dW_i^{\mathbb{Q}}(t) = dW_i(t) + \frac{1}{m(t)} d\langle W_i(t), m(t)\rangle = \lambda_i(t)\, dt + dW_i(t) \tag{2.25}$$

where $W_i^{\mathbb{Q}}(t)$ is a standard Brownian motion under the risk-neutral measure $\mathbb{Q}$.

The dynamics of $S(t)$, $\sigma_i(t)$, and $v_{TK}(t)$ under the risk-neutral measure are thus given by the following equations:

$$\begin{aligned}
\frac{dS}{S} &= rdt + \sum_{i=1}^{\infty}\sum_{j=0}^{\infty}\kappa_{ij}\sigma_i dW_j^{\mathbb{Q}} && (2.26)\\
d\sigma_i &= (\gamma_i - \lambda_i\delta_i)\,dt + \delta_i dW_i^{\mathbb{Q}} \qquad i = 1,\,2,\ldots && (2.27)\\
d_t v_{TK} &= \left(\alpha_{TK} - \beta_{TK}\sum_{i=0}^{\infty}\zeta_{TKi}\lambda_i\right)dt + \beta_{TK}\sum_{i=0}^{\infty}\zeta_{TKi}dW_i^{\mathbb{Q}} && (2.28)
\end{aligned}$$

and the dynamics of $v_{tS} = \sqrt{\sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right)^2}$ are:

$$\begin{aligned}
dv_{tS} = &\left[\frac{\partial v}{\partial T} + rS\frac{\partial v}{\partial K} + S\sum_{j=0}^{\infty}\frac{\partial}{\partial K}\left(\beta_{tS}\zeta_{tSj}\right)\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i + \frac{1}{2}\frac{\partial^2 v}{\partial K^2}S^2\sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right)^2\right.\\
&\left.+\alpha_{tS} - \beta_{tS}\sum_{j=0}^{\infty}\zeta_{tSj}\lambda_j\right]dt + \sum_{j=0}^{\infty}\left(\beta_{tS}\zeta_{tSj} + \frac{\partial v}{\partial K}S\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i\right)dW_j^{\mathbb{Q}} \qquad (2.29)
\end{aligned}$$

The risk-neutral dynamics of v_{tS} can be used to price volatility derivatives such as volatility swaps and swaptions. An investor who enters a long position in a volatility swap at time $t = 0$, with time to maturity T, receives the following payoff at maturity:

$$\frac{1}{T}\int_0^T u(x)\,dx - K$$

where $u(t)$ represents the measure of volatility being used for the swap, and the constant K is the fixed payment. In order that the value of the swap is equal to zero at $t = 0$, K must be set equal to the expected value of the floating part under the risk-neutral measure:

$$K = E_0^{\mathbb{Q}}\left(\frac{1}{T}\int_0^T u(x)\,dx\right) = \frac{1}{T}\int_0^T E_0^{\mathbb{Q}}\left[u(x)\right]dx \qquad (2.30)$$

where the subscript zero in the expectation operator denotes that the expectation is taken with respect to the information at time zero. If $u = v_{tS}$ then u represents the instantaneous volatility of the stock price process. If $u = v_{tS}^2$ then the security is a variance swap. Suppose that the dynamics of $\omega = v_{tS}^2$ under the risk-neutral measure are given by the following equation:

$$d\omega_t = \theta\left(\rho - \omega_t\right)dt + \sqrt{\omega_t - \nu}\sum_{j=0}^{\infty}\psi_j dW_j^{\mathbb{Q}} \qquad (2.31)$$

where $\theta,\ \rho,\ \nu$, and ψ_j are constants $\forall\ j,\ \nu > 0$, and ω_t denotes $\omega(t)$. Under the above specification the process ω_t has a reflecting barrier at ν, so that $\omega_t \geq \nu > 0\ \forall\ t$. The expectation in equation (2.30) is equal to:

$$E_0^{\mathbb{Q}}\left(\omega_t\right) = e^{-\theta t}\left(\omega_0 - \rho\right) + \rho$$

Hence, the fixed payment in a variance swap, K_ω, is:

$$K_\omega = \frac{1}{T\theta}\left(1 - e^{-\theta T}\right)(\omega_0 - \rho) + \rho$$

Call and put variance swaptions have the following payoff at maturity T, respectively:[16]

$$\max\left(\frac{1}{T}\int_0^T \omega_t dt - K_\omega, 0\right) \text{ call variance swaption}$$
$$\max\left(K_\omega - \frac{1}{T}\int_0^T \omega_t dt, 0\right) \text{ put variance swaption}$$

We define the following variable:

$$V_t = \int_0^t \omega_x dx$$

Hence,

$$dV_t = \omega_t dt$$

The price C_0^ω of a call variance swaption at time $t = 0$ is equal to:

$$\begin{aligned} C_0^\omega &= e^{-rT} E_0^Q\left[\max\left(\frac{V_T}{T} - K_\omega, 0\right)\right] \\ &= \frac{e^{-rT}}{T} E_0^Q\left(V_T I_{V_T \geq TK_\omega}\right) - e^{-rT} K_\omega P_0^Q\left(V_T \geq TK_\omega\right) \\ &= \frac{e^{-rT}}{T}\int_{TK_\omega}^{\infty} P_0^Q\left(V_T \geq x\right) dx \end{aligned} \tag{2.32}$$

where $I_{V_T \geq TK_\omega}$ is the indicator function and $P_0^Q\left(V_T \geq TK_\omega\right)$ denotes the probability that $V_T \geq TK_\omega$ at time $t = 0$ under the risk-neutral measure. To derive equation (2.32) we use the following relationship between the probability density function $f(x)$ and the distribution function $F(x)$ of a random variable x, obtained by integrating by parts:

$$\int_0^z x f(x)\, dx = zF(z) - \int_0^z F(x)\, dx$$

$P_0^Q\left(V_T \leq x\right)$ can be obtained in terms of the inverse Fourier transform of the characteristic function of V_T under the risk-neutral measure, denoted by $f(V_t, \omega_t, t, T; \varphi)$ (Kendall, 1987):[17]

$$P_0^Q\left(V_T \leq x\right) = \frac{1}{2} - \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\varphi x} f\left(V_0, \omega_0, 0, T; \varphi\right)}{i\varphi}\right] d\varphi \tag{2.33}$$

[16] Variance swaptions may also be defined in terms of the discrete-time variance of the stock price.

[17] If V_t were a traded asset, we could use a change of measure to convert $e^{-rT} E_0^Q\left(V_T I_{V_T \geq TK_\omega}\right)$ into V_t multiplied by a probability under the appropriate measure.

The price G_0^ω of a put variance swaption at time $t = 0$ is equal to:

$$\begin{aligned} G_0^\omega &= e^{-rT} E_0^{\mathbb{Q}} \left[\max \left(K_\omega - \frac{V_T}{T}, 0 \right) \right] \\ &= \frac{e^{-rT}}{T} \left[K_\omega T P_0^{\mathbb{Q}} \left(V_T \leq T K_\omega \right) - E_0^{\mathbb{Q}} \left(V_T I_{V_T \leq T K_\omega} \right) \right] \\ &= \frac{e^{-rT}}{T} \int_0^{TK_\omega} P_0^{\mathbb{Q}} \left(V_T \leq x \right) dx \end{aligned}$$

The characteristic function of V_T is the solution to Kolmogorov's backward differential equation:

$$0 = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial V} \omega_t + \frac{\partial f}{\partial \omega} \theta \left(\rho - \omega_t \right) + \frac{1}{2} \frac{\partial^2 f}{\partial \omega^2} \xi \left(\omega_t - \nu \right) \tag{2.34}$$

where $\xi = \sum_{i=0}^{\infty} \psi_i^2$, subject to the boundary condition

$$f \left(V_T, \omega_T, T, T; \varphi \right) = e^{i \varphi V_T}$$

The characteristic function solution has the following form:

$$f \left(V_t, \omega_t, t, T; \varphi \right) = e^{D_t + M_t \omega_t + i \varphi V_t}$$

Substituting this solution into (2.34) we obtain

$$0 = \frac{dD}{dt} + \frac{dM}{dt} \omega_t + i \varphi \omega_t + M_t \theta \left(\rho - \omega_t \right) + M_t^2 \frac{\xi}{2} \left(\omega_t - \nu \right)$$

Hence D_t and M_t are the solutions to the following system of ordinary differential equations, subject to the boundary conditions $D_T = 0$ and $M_T = 0$:

$$\frac{dD}{dt} = -M_t \theta \rho + M_t^2 \frac{\xi \nu}{2} \qquad \qquad \frac{dM}{dt} = -i \varphi + M_t \theta - \frac{1}{2} M_t^2 \xi$$

The solutions are:

$$\begin{aligned} D_t &= - \left[N \left(\Gamma + \Delta \right) + \Lambda \left(\Gamma^2 - \Delta^2 \right) \right] \ln \left(1 - B e^{-A \xi t} \right) \\ & \quad - \Lambda \frac{\left(\Gamma + \Delta \right)^2}{1 - B e^{-A \xi t}} + \left(\Lambda \Delta^2 - N \Delta \right) A \xi t + c \\ M_t &= \Delta \frac{1 - e^{A \xi (T - t)}}{1 - e^{-A \xi t} B} \end{aligned}$$

where

$$A = \sqrt{\frac{\theta^2}{\xi^2} - \frac{2 i \varphi}{\xi}} \qquad B = \frac{A + \frac{\theta}{\xi}}{\frac{\theta}{\xi} - A} e^{A \xi T} \qquad \Gamma = A - \frac{\theta}{\xi}$$

$$\Delta = \frac{\theta}{\xi} + A \qquad \Lambda = \frac{\nu}{2A} \qquad N = \frac{\theta \rho}{A \xi}$$

and c is a constant determined by the boundary condition for D_t.

We can also obtain a closed-form solution for the price of a derivative with the following payoff at maturity T:

$$\max\left(e^{(L-\omega_T)T}-K,0\right) \tag{2.35}$$

where L and K are constants.

The price Φ_0 at time $t=0$ of the derivative with payoff (2.35) is equal to:

$$\begin{aligned}\Phi_0 &= e^{-rT}E_0^{\mathbb{Q}}\left[\max\left(e^{(L-\omega_T)T}-K,0\right)\right]\\ &= e^{(L-r)T}E_0^{\mathbb{Q}}\left(e^{-T\omega_T}I_{\omega_T\leq L-\frac{1}{T}\ln K}\right)-e^{-rT}KP_0^{\mathbb{Q}}\left(\omega_T\leq L-\frac{1}{T}\ln K\right)\end{aligned}$$

The probability in the above equation is obtained from the following inverse Fourier transform:

$$P_0^{\mathbb{Q}}\left(\omega_T\leq L-\frac{1}{T}\ln K\right)=\frac{1}{2}-\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\varphi\left(L-\frac{1}{T}\ln K\right)}f\left(\omega_0,0,T;\varphi\right)}{i\varphi}\right]d\varphi$$

where $f\left(\omega_t,t,T;\varphi\right)$ is the characteristic function of ω_T.

$E_0^{\mathbb{Q}}\left(e^{-T\omega_T}I_{\omega_T\leq L-\frac{1}{T}\ln K}\right)$ is obtained using the following lemma:

Lemma 2.3 *Let y denote a random variable and $f\left(\varphi\right)$ its characteristic function, $f\left(\varphi\right)=E\left(e^{i\varphi y}\right)$. $E\left(e^{\alpha y}I_{y\geq K}\right)$, where α and K are constants, is equal to:*

$$E\left(e^{\alpha y}I_{y\geq K}\right)=\frac{f\left(-i\alpha\right)}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{f\left(\varphi-i\alpha\right)e^{-i\varphi K}}{i\varphi}\right]d\varphi \tag{2.36}$$

PROOF. In appendix B.

Thus,

$$\begin{aligned}E_0^{\mathbb{Q}}\left(e^{-T\omega_T}I_{\omega_T\leq L-\frac{1}{T}\ln K}\right) &= \frac{f\left(\omega_0,0,T;iT\right)}{2}\\ &\quad -\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\varphi\left(L-\frac{1}{T}\ln K\right)}f\left(\omega_0,0,T;\varphi+iT\right)}{i\varphi}\right]d\varphi\end{aligned}$$

The characteristic function of ω_T satisfies Kolmogorov's backward equation:

$$0=\frac{\partial f}{\partial t}+\frac{\partial f}{\partial \omega}\theta\left(\rho-\omega_t\right)+\frac{1}{2}\frac{\partial^2 f}{\partial \omega^2}\xi\left(\omega_t-\nu\right) \tag{2.37}$$

subject to the boundary condition

$$f\left(\omega_T,T,T;\varphi\right)=e^{i\varphi\omega_T}$$

The characteristic function solution has the following form:

$$f\left(\omega_t,t,T;\varphi\right)=e^{D_t+M_t\omega_t+i\varphi\omega_t}$$

Substituting this solution into equation (2.37) we obtain:

$$0 = \frac{dD}{dt} + \frac{dM}{dt}\omega_t + (i\varphi + M_t)\,\theta\,(\rho - \omega_t) + (i\varphi + M_t)^2\,\frac{\xi}{2}\,(\omega_t - \nu) \qquad (2.38)$$

subject to the boundary conditions $D_T = 0$ and $M_T = 0$.

Equation (2.38) reduces to the following system of ordinary differential equations:

$$\begin{aligned} \frac{dD}{dt} &= -(i\varphi + M_t)\,\theta\rho + (i\varphi + M_t)^2\,\frac{\xi\nu}{2} \\ \frac{dM}{dt} &= (i\varphi + M_t)\,\theta - (i\varphi + M_t)^2\,\frac{\xi}{2} \end{aligned}$$

The solution is:

$$\begin{aligned} D_t &= \varphi^2 t\,(\xi\nu - 2\Lambda) + \frac{4\Lambda\theta}{\xi^2\,(1 - Ae^{\theta t})} \\ &\quad + \left[\frac{\Lambda}{\theta}\,(B^2 + \varphi^2) - \frac{2}{\xi}\,(\rho\theta - i\varphi\xi\nu)\right] \ln\left(1 - Ae^{\theta t}\right) + c \\ M_t &= i\varphi\frac{e^{\theta(t-T)} - 1}{1 - e^{\theta t}A} \end{aligned}$$

where

$$A = \frac{i\varphi e^{-\theta T}}{i\varphi - \frac{2\theta}{\xi}} \qquad B = i\varphi - \frac{2\theta}{\xi} \qquad \Lambda = \frac{\xi\nu}{2}$$

and c is a constant determined by the boundary condition for D_t.

2.4 Pricing and hedging contingent claims

In incomplete markets there exist an infinite number of equivalent martingale measures (EMMs), with each specification of the volatility risk premium (or premia, in the case of more than one volatility or factor) corresponding to a particular EMM. When the volatility risk premium is not zero, contingent claims cannot be perfectly hedged using only stocks and the riskless asset, in general. Options need to be added to the hedging portfolio in order to attain a perfect hedge.

In the models of Schonbucher (1999) and Ledoit *et al.* (2002) all derivatives can be priced as a function of the stock price and the IV of at-the-money options with zero time to maturity, which in practice can be approximated using close-to-maturity at-the-money options. Our model, however, implies that the IV processes of all plain vanilla European options are needed to price other derivatives, and hedging portfolios for exotic contingent claims must contain European options of all maturity dates and exercise prices, in general. Although the construction of such a hedging portfolio is unfeasible in practice, it may be possible to attain a near-perfect hedge using only a small number of options which are highly correlated with the contingent claim, since the weight of each option in the hedging portfolio is related to the correlation between the option and the contingent claim.

This result is very different from the well-known fact that any contingent claim can be perfectly hedged using any one option and the stock in models with a single stochastic volatility process, but is similar to the situation in random field models of the term structure of interest rates, where it may be possible to attain a near-perfect hedge for an interest rate contingent claim using only those bonds that are most correlated with this derivative asset. This feature of random field term structure models is in fact one of their advantages with respect to the HJM framework, because it is consistent with market hedging practice, whereas n-factor HJM models, which predict that an interest rate contingent claim can be hedged using any n bonds, are not. The correlation between interest rate contingent claims and bonds of different maturities depends on the maturity characteristics of the derivative asset and, as a result, it is usually easy to select the appropriate bonds for hedging purposes. In the case of stock options, which have an exercise price as well as a maturity dimension, however, it is more difficult to ascertain which options to use in order to hedge an exotic derivative.

In section (2.2) we showed that under the assumptions of our model derivative securities can be priced as a function of the stock price and all implied volatilities. One can estimate the dynamics of the IV surface using a small number of parameters by assuming, for example, that the drift and volatility of IV are smooth functions of the exercise price and the time to maturity. In general, the price $U(t)$ of a contingent claim satisfies the following stochastic partial differential equation, which can be solved numerically, if no closed-form solution exists:

$$\begin{aligned} dU &= \left[\frac{\partial U}{\partial t} + \frac{\partial U}{\partial S} rS + \frac{1}{2}\frac{\partial^2 U}{\partial S^2} S^2 \sum_{j=0}^{\infty}\left(\sum_{i=1}^{\infty}\sigma_i \kappa_{ij}\right)^2 \right. \\ &\quad + \int_T \int_K \frac{\partial U}{\partial v_{TK}}\left(\alpha_{TK} - \beta_{TK}\sum_{j=0}^{\infty}\zeta_{TKj}\lambda_j\right) dT dK \\ &\quad + \frac{1}{2}\int_{T_1}\int_{K_1}\int_{T_2}\int_{K_2} \frac{\partial^2 U}{\partial v_{T_1K_1}\partial v_{T_2K_2}} \beta_{T_1K_1}\beta_{T_2K_2}\sum_{i=0}^{\infty}\zeta_{T_1K_1i}\zeta_{T_2K_2i} dT_1 dK_1 dT_2 dK_2 \\ &\quad \left. + \int_T\int_K \frac{\partial^2 U}{\partial S \partial v_{TK}} S\beta_{TK}\sum_{j=0}^{\infty}\zeta_{TKj}\sum_{i=1}^{\infty}\kappa_{ij}\sigma_i dT dK \right] dt \\ &\quad + \sum_{j=0}^{\infty}\left(\frac{\partial U}{\partial S} S \sum_{i=1}^{\infty}\kappa_{ij}\sigma_i + \int_T\int_K \frac{\partial U}{\partial v_{TK}}\beta_{TK}\zeta_{TKj} dT dK\right) dW_j^{\mathbb{Q}} \end{aligned} \tag{2.39}$$

2.5 Conclusion

In this chapter we develop a model for the Black-Scholes IV of plain-vanilla European stock options where the stochastic shocks driving IV form a random field with three parameters: current time, the maturity date and the exercise price of the corresponding option. The drift of the IV process is endogenously determined under the no-arbitrage condition on option prices. Using this drift restriction we show that as the time to maturity tends to zero and the exercise price tends to the current stock price, IV converges to a measure of the instantaneous volatility of the underlying asset. Under the assumptions of the model, the entire continuum of plain-vanilla European options for all maturity dates and exercise prices is required to complete the market.

Using the dynamics of IV, illiquid and exotic derivatives can be priced, taking the prices of liquid plain-vanilla European options as given. Our model offers greater flexibility than the diffusion IV models of Schonbucher (1999) and Ledoit *et al.* (2002) by having different, but correlated, stochastic shocks for each maturity date and exercise price. By modelling IV as a random field we can fit the initial IV surface exactly, in contrast to previous stochastic volatility models. Fitting initial plain vanilla option prices exactly is an advantage if these options are to be used to hedge more complex derivatives.

There exists much scope for further research in this area, such as the empirical implementation of the model and comparisons of its performance against other stochastic volatility models. Furthermore, the model can be extended to include jumps in returns and/or volatility,[18] and stochastic interest rates, and can also be applied to the IV of foreign currency and interest rate derivatives.

[18] Eraker, Johannes, and Polson (2003) provide some recent evidence of the existence of jumps in both returns and volatility.

Chapter 3

Random-field models of the term structure of interest rates with stochastic volatility

3.1 Introduction

Models of the term structure of interest rates can be classified according to the variable being modelled: spot rate models, instantaneous forward rate models, and LIBOR or market models. Spot rate models were the first to be studied (Merton, 1973; Vasicek, 1977; Cox, Ingersoll, and Ross, 1985). Initial spot rate models were one-factor models. In such models all forward rates are perfectly correlated. In practice, however, the correlation coefficient between the returns on bonds of different maturities is less than one. Having a model that can handle imperfect correlation is especially important when pricing derivatives which depend on multiple rates with different maturities. Spot rate models were extended to include multiple factors by adding more sources of stochastic shocks, such as stochastic volatility and/or imperfectly correlated long-maturity rates. For example, Brennan and Schwartz (1979) present a two-factor model for the spot rate and the long-term rate, the latter being defined as the yield on a consol bond which pays coupons continuously. These two variables are assumed to follow a joint diffusion process. Rebonato and Cooper (1996) find that the forward-rate correlation structure generated by two-factor models is also inconsistent with the data, because the correlation does not decrease fast enough as the difference in maturities rises.

The next generation of term structure models, introduced by Ho and Lee (1986), were constructed to fit a given initial forward rate curve exactly. Ho and Lee (1986) develop a discrete-time model where the term structure follows a binomial lattice. Heath, Jarrow, and Morton (HJM) (1992) extend this approach to continuous time. Models of this type have the advantage of being preference-free, and also offer flexibility with respect to the number of factors driving the forward rate curve. In a HJM model with n factors, an interest rate contingent claim can be hedged using any set of n bonds. This, however, is inconsistent with market practice, where contingent claims tend to be hedged with bonds of similar maturity.

In recent years there has been increasing interest in random field (or string) models of the term structure of interest rates. In this class of models, introduced by Kennedy (1994), each instantaneous forward rate is driven by a different stochastic shock process, that is correlated with those driving other forward rates. As a result, random field models are infinite-factor models.[1] Models of this type can be easier to calibrate than HJM models. For example, Santa-Clara and Sornette (2001) calibrate two random field forward rate models which are completely specified by only three parameters. Furthermore, random field models have the advantage that they generate hedging portfolios for interest rate contingent claims that are more consistent with market practice than HJM models.[2]

The simplest random field term structure model assumes that interest rate volatility is deterministic[3] and that interest rate innovations are generated by a Brownian random field. Kennedy (1994) studies a Gaussian random field model for forward rates, and derives the restriction on the drift of the instantaneous forward rate process imposed by the no-arbitrage condition. He also obtains a closed-form solution for the price of interest rate caplets. In a subsequent paper (Kennedy, 1997) he shows that when forward rates are generated by a random field and are Gaussian, Markov, and stationary, the forward rate volatility and correlation functions depend on three parameters only. Chu (1996) prices a number of interest rate derivatives using the partial differential equations approach in a random field model for the yield curve, where the random

[1]Santa-Clara and Sornette (2001) argue that there is a distinction between infinite-factor HJM models and random field models. Infinite-factor HJM models are in fact random field models when each forward rate is driven by a combination of the factors that is linearly independent of those driving other forward rates – see Filipovic (2000) for example. The class of random field term structure models, however, is wider than the class of infinite-factor HJM models, since the random field innovations need not necessarily be Gaussian.

[2]In general, bonds of all maturities are needed in order to hedge an interest rate contingent claim in a random field model. If markets are incomplete due to stochastic volatility, options are also needed to construct a perfect hedge. Nevertheless, it may be possible to hedge most of the variability in a contingent claim using one or a few bonds of similar maturity.

[3]Similarly to chapter 2, forward rate volatility is said to be deterministic if it can be written as a function of interest rates, time, and maturity only. The term stochastic volatility is reserved for the case where volatility introduces additional risk factors in the economy.

innovations are generated by a Brownian sheet and the covariance function of bond returns is a deterministic function of time. Goldstein (2000) extends Kennedy's (1994) drift restriction to non-Gaussian random fields and prices bond options using Heston's (1993) inverse Fourier transform method. He assumes that forward rate volatility and correlation are deterministic functions of the spot rate and the time to maturity. Santa-Clara and Sornette (2001) calibrate forward rate models using two types of random field shocks: an Ornstein-Uhlenbeck random field and a subexponential correlation random field. They assume that forward rate volatility is a deterministic function of the time to maturity.

Numerous empirical studies suggest that the volatility of interest rates is stochastic – Ball and Torous (1999), Ahn *et al.* (2003), Dai and Singleton (2003) are some recent examples. Amin and Morton (1994) test the HJM model with time- and level-dependent volatility using Eurodollar futures and options data, and find systematic strike-price and time-to-maturity biases. The implied volatility smiles exhibited by interest rate derivative prices also offer indirect evidence that interest rate volatility is stochastic. Jarrow, Li, and Zhao (2003) find that the implied volatility of interest rate caplets[4] forms an asymmetric smile when plotted against the cap strike rate. Similarly, Rebonato (2003) presents implied volatility smiles for the swaptions market. Other recent empirical studies test for the presence of stochastic volatility by examining whether the prices of interest rate derivatives are spanned by bond prices. Collin-Dufresne and Goldstein (2002), Collin-Dufresne, Goldstein, and Jones (2003), and Heidari and Wu (2003) offer empirical evidence that interest rate derivatives are not redundant securities and that interest rate volatility risk cannot be hedged by a portfolio consisting of bonds only.[5]

Spot rate models with stochastic volatility have been widely studied (see for example Fong and Vasicek (1991), Longstaff and Schwartz (1992), and Chen (1996)). Stochastic volatility HJM models have received less attention, due to the difficulties they present in the pricing of interest rate contingent claims and the estimation of model parameters. Chiarella and Kwon (1999) consider a stochastic volatility HJM model which can be transformed into a Markovian system. They derive closed-form solutions for discount bond prices in terms of the state variables in the Markovian system and value bond options numerically. Andreasen, Collin-Dufresne, and Shi (1998) construct a stochastic volatility HJM model with the additional assumption that there exists a market in futures

[4]The implied volatility of interest rate caplets is calculated using the Black (1976) model.

[5]In contrast to these studies, Fan, Gupta, and Ritchken (2002) find that swaptions can be adequately hedged using bonds. Collin-Dufresne and Goldstein (2003), however, offer an explanation for this result that is consistent with the presence of unspanned stochastic volatility.

on the square of bond price volatility, or that such futures can be perfectly replicated using traded assets. They show that the latter assumption is satisfied if there exist futures on yields, which are available in practice. They price interest rate derivatives using a non-recombining trinomial tree that fits initial bond and volatility futures prices. The prices of volatility futures serve to obtain the volatility risk premium.

In the context of a multifactor term structure model, Duffie and Kan (1996) derive the necessary and sufficient conditions for bond prices to be exponential affine functions of the factors. One of these conditions is that the spot rate is an affine function of the factors. Collin-Dufresne and Goldstein (2002) provide some examples of affine term structure models that generate unspanned stochastic volatility. Duffie, Pan, and Singleton (2000) derive closed-form solutions for derivative prices in this framework, in terms of Fourier transforms. Collin-Dufresne and Goldstein (2003) extend this class of models to include an infinite number of factors. They assume that the continuum of bond prices and bond-price volatility are the state variables in this setting. Bond prices are driven by a Brownian random field and bond price volatility by a standard Brownian motion.

In this chapter we construct random field models of instantaneous forward, LIBOR, and swap rates with stochastic volatility, and derive closed-form solutions for the prices of a number of interest rate derivatives. The rest of this chapter is organized as follows. Section 3.2 presents a random field model for instantaneous forward rates. Forward rate volatility is assumed to be the product of a deterministic function of the time to maturity and a maturity-independent stochastic process driven by a standard Brownian motion. We allow for non-zero correlation between this Brownian motion and the random field driving forward rates. The separability of the volatility function enables us to obtain closed-form solutions for the prices of bond options, interest rate caplets, and interest rate spread options, using Heston's (1993) inverse Fourier transform method. In this model markets are incomplete due to stochastic forward rate volatility. In Section 3.3, after a brief introduction to market models, we present stochastic volatility random field models for forward LIBOR and swap rates with separable volatility structure, and price LIBOR caplets and swaptions. Section 3.4 concludes.

3.2 A random-field model for the forward rate with stochastic volatility

3.2.1 The model under the risk-neutral measure

Uncertainty is modelled as a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$, where $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ is the $\mathbb{Q}$-augmentation of the natural filtration

$$\mathcal{F}_t = \sigma\left(Z\left(u, s\right), W\left(u\right), 0 \leq u \leq t, s \geq t\right)$$

generated by the random field $Z(t, s)$ and a standard Brownian motion $W(t)$, where t denotes current time and s the maturity date, $t \geq 0,\ s \geq t$. We assume that $\mathbb{F}$ satisfies the *usual conditions*. $\mathbb{Q}$ represents the risk-neutral measure.

There exist zero-coupon risk-free bonds for all maturities $s \geq t$. The time-t price of a zero-coupon risk-free bond which pays one currency unit at time $s \geq t$ is denoted by $P(t, s)$. The instantaneous forward rate $f(t, s)$ at time t for maturity date s is defined as the interest rate that can be contracted at time t to be received on an instantaneous riskless loan at time s:

$$f(t, s) = -\frac{\partial \ln P(t, s)}{\partial s} \tag{3.1}$$

and

$$P(t, s) = e^{-\int_t^s f(t,x)dx} \tag{3.2}$$

The spot rate $r(t)$ is defined as the instantaneous interest rate that is received on a riskless loan at time t. Thus $r(t) = f(t, t)$. There exists a bank account with price process $B(t)$ adapted to the filtration $\{\mathcal{F}_t\}$. We assume that $B(0) = 1$. The value of the bank account at time t is:

$$B(t) = e^{\int_0^t r(x)dx} \tag{3.3}$$

Hence,

$$dB(t) = B(t)\, r(t)\, dt \tag{3.4}$$

The dynamics of the instantaneous forward rate under the risk-neutral measure are assumed to be as follows:

$$d_t f(t, s) = \alpha(t, s)\, dt + \sigma(t, s)\, d_t Z(t, s) \tag{3.5}$$

or, in integral form,

$$f(t, s) = f(0, s) + \int_0^t \alpha(x, s)\, dx + \int_0^t \sigma(x, s)\, d_x Z(x, s) \tag{3.6}$$

where the subscript t in the differential operator indicates that the increment is obtained with respect to time t.

ASSUMPTION I *The random field $Z(t,s)$ satisfies the following conditions:*

(i) $Z(t,s)$ *is continuous in* t *and* s.

(ii) $E[d_t Z(t,s)] = 0$.

(iii) $var[d_t Z(t,s)] = dt$.

(iv) $cor[d_t Z(t,s_1), d_t Z(t,s_2)] = c(s_1 - s_2) = c(s_2 - s_1)$.

(v) $d_t Z(t_1,s_1)$ *and* $d_t Z(t_2,s_2)$ *are independent* $\forall \;\; t_1 \neq t_2$.

Under the above conditions the correlation structure of the random field $Z(t,s)$ does not depend on t. The instantaneous volatility of the forward rate, $\sigma(t,s)$, is assumed to be of the following form:

$$\sigma(t,s) = \sqrt{\omega(t)}g(s-t) \tag{3.7}$$

where $g(u)$ is a deterministic function of u and the dynamics of $\omega(t)$ are:

$$d\omega(t) = \beta[\gamma - \omega(t)]\,dt + \delta\sqrt{\omega(t)}dW(t) \tag{3.8}$$

where $W(t)$ is a standard Brownian motion, and β, γ, and δ are constants.

ASSUMPTION II *(i)* $cor[d_t Z(t,s), dW(t)] = \rho(s-t)$.

(ii) $d_t Z(t_1,s)$ *and* $dW(t_2)$ *are independent for* $t_1 \neq t_2$.

The separability of the volatility function $\sigma(t,s)$ into a deterministic function of the time to maturity and a maturity-independent stochastic process implies that the volatilities of different maturities are perfectly correlated. Although this reduces the flexibility of the model, it allows us to derive closed-form solutions for the prices of a number of interest rate derivatives. Greater flexibility could be achieved by specifying a more general volatility function, such as that proposed by Goldstein (2000):

$$d\sigma^2(t,s) = \kappa(s-t)\left[\theta(s-t) - \sigma^2(t,s)\right]dt + \nu(s-t)\,\sigma(t,s)\,d_t Z_\sigma(t,s)$$

where volatility is driven by random field innovations which can be correlated with those of forward rates:

$$cor[d_t Z_\sigma(t,s_1), d_t Z(t,s_2)] = \xi(t,s_1,s_2)$$

The above model, however, does not yield closed-form solutions for the prices of interest rate derivatives and is more difficult to calibrate.

Proposition 3.1 *Under assumptions (I), (II), and the no-arbitrage condition on bond prices, the drift of the forward rate is equal to:*

$$\alpha(t,s) = \sigma(t,s)\int_t^s \sigma(t,y)\,c(s-y)\,dy \tag{3.9}$$

PROOF. In appendix C.

Under the forward rate volatility specification (3.7), the risk-neutral dynamics of forward rates and bond prices are:

$$\begin{aligned} d_t f(t,s) &= \omega(t)\, g(s-t) \int_t^s g(y-t)\, c(s-y)\, dy dt + \sqrt{\omega(t)} g(s-t)\, d_t Z(t,s) \\ \frac{d_t P(t,s)}{P(t,s)} &= r(t)\, dt - \sqrt{\omega(t)} \int_t^s g(y-t)\, d_t Z(t,y)\, dy \end{aligned}$$

Collin-Dufresne and Goldstein (2003) study a random field term structure model which is a special case of our model with $\rho(s) = 0\ \forall\ s$ and $Z(t,s)$ being Gaussian. They also allow the correlation structure of the random field to be time-varying, and consider the case where this correlation follows a two-state Markov chain. Kimmel (2004) extends their model to include multiple latent variables that follow a joint diffusion process.

Under the risk-neutral measure, discounted asset price processes are martingales. The numeraire used for discounting in this case is the bank account process $B(t)$. If instead of the bank account we use the price of a zero-coupon risk-free bond with maturity T as the numeraire, then there exists a measure, the T-forward measure $\mathbb{Q}^T$, under which asset prices discounted by $P(t,T)$ are martingales (Geman, El Karoui, and Rochet, 1995).

Proposition 3.2 *Under the T-forward measure the random field $Z^T(t,s)$ defined by the equation below satisfies assumption (I):*

$$d_t Z^T(t,s) = d_t Z(t,s) + \int_t^T \sigma(t,y)\, c(s-y)\, dy dt \tag{3.10}$$

PROOF. In appendix D.

Let

$$\frac{dm(t)}{m(t)} = \int_t^T \sigma(t,y)\, d_t Z(t,y)\, dy \tag{3.11}$$

Then $W^T(t)$ defined by the equation below is a Brownian motion under the T-forward measure:

$$\begin{aligned} dW^T(t) &= dW(t) + \frac{1}{m(t)} d\langle W(t), m(t)\rangle \\ &= dW(t) + \int_t^T \sigma(t,y)\, \rho(y-t)\, dy dt \end{aligned} \tag{3.12}$$

$Z^T(t,s)$ and $W^T(t)$ also satisfy assumption (II).

Under the T-forward measure the forward rate with maturity T is a martingale:

$$d_t f(t,T) = \sigma(t,T)\, d_t Z^T(t,T)$$

The T-forward measure is useful for pricing interest rate derivatives. Let $X(t)$ denote the price of an asset at time t. Under the T-forward measure, $X(t)$ discounted by $P(t,T)$ is a martingale:

$$\frac{X(t)}{P(t,T)} = E_t^{\mathbb{Q}^T}\left[\frac{X(v)}{P(v,T)}\right] \quad \text{for } t \leq v \leq T$$

Under the risk-neutral measure, $X(t)$ discounted by $B(t)$ is a martingale:

$$\frac{X(t)}{B(t)} = E_t^{\mathbb{Q}}\left[\frac{X(v)}{B(v)}\right] \quad \text{for } t \leq v$$

From the previous two equations we obtain:

$$B(t)\, E_t^{\mathbb{Q}}\left[\frac{X(v)}{B(v)}\right] = P(t,T)\, E_t^{\mathbb{Q}^T}\left[\frac{X(v)}{P(v,T)}\right] \text{ for } t \leq v \leq T \tag{3.13}$$

If $v = T$ then

$$B(t)\, E_t^{\mathbb{Q}}\left[\frac{X(T)}{B(T)}\right] = P(t,T)\, E_t^{\mathbb{Q}^T}\left[X(T)\right] \tag{3.14}$$

3.2.2 Pricing bond options

Let $C(t,m,T,K)$ denote the time-t price of a European call option with maturity $m \geq t$ and exercise price K, written on a bond $P(t,T)$ with maturity $T \geq m$. The payoff of the option at maturity m is:

$$C(m,m,T,K) = \max\left[P(m,T) - K, 0\right] \tag{3.15}$$

Under the risk-neutral measure, the discounted price process of the call option is a martingale. Hence,

$$\begin{aligned} C(t,m,T,K) &= E_t^{\mathbb{Q}}\left\{e^{-\int_t^m r(x)dx} I_{P(m,T)\geq K}\left[P(m,T) - K\right]\right\} \\ &= B(t)\, E_t^{\mathbb{Q}}\left[\frac{P(m,T)}{B(m)} I_{P(m,T)\geq K}\right] - B(t)\, K E_t^{\mathbb{Q}}\left[\frac{I_{P(m,T)\geq K}}{B(m)}\right] \end{aligned}$$

Using equations (3.13) and (3.14), the above equation becomes:

$$C(t,m,T,K) = P(t,T)\, P_t^{\mathbb{Q}^T}\left[P(m,T) \geq K\right] - K P(t,m)\, P_t^{\mathbb{Q}^m}\left[P(m,T) \geq K\right]$$

where $P_t^{\mathbb{Q}^T}\left[P(m,T) \geq K\right]$ denotes the probability at time t that $P(m,T) \geq K$ under measure $\mathbb{Q}^T$.

Let $\psi(v,t;\varphi)$ denote the characteristic function of $\ln P(m,T)$ under the v-forward measure:

$$\psi(v,t;\varphi) = E_t^{\mathbb{Q}^v}\left[e^{i\varphi \ln P(m,T)}\right]$$

Using the inverse Fourier transform of the characteristic function, the bond option price is equal to:

$$\begin{aligned} C(t,m,T,K) &= P(t,T)\left\{\frac{1}{2}+\frac{1}{\pi}\int_0^{\infty}\mathrm{Re}\left[\frac{e^{-i\varphi\ln K}\psi(T,t;\varphi)}{i\varphi}\right]d\varphi\right\} \\ &\quad -KP(t,m)\left\{\frac{1}{2}+\frac{1}{\pi}\int_0^{\infty}\mathrm{Re}\left[\frac{e^{-i\varphi\ln K}\psi(m,t;\varphi)}{i\varphi}\right]d\varphi\right\} \end{aligned}$$

The dynamics of $\ln P(t,T)$, $r(t)$, and $\omega(t)$ under the v-forward measure are:

$$\begin{aligned} d_t\ln P(t,T) &= \left[r(t)-\frac{1}{2}\omega(t)\int_t^T\int_t^T g(u-t)\,g(y-t)\,c(u-y)\,dudy\right]dt \\ &\quad +\omega(t)\int_t^T\int_t^v g(u-t)\,g(y-t)\,c(u-y)\,dudydt \\ &\quad -\sqrt{\omega(t)}\int_t^T g(u-t)\,d_tZ^v(t,u)\,du \\ dr(t) &= \left[\left.\frac{\partial f(t,s)}{\partial s}\right|_{s=t}-\omega(t)\,g(0)\int_t^v g(y-t)\,c(t-y)\,dy\right]dt \\ &\quad +\sqrt{\omega(t)}g(0)\,dZ^v(t,t) \\ d\omega(t) &= \left\{\beta[\gamma-\omega(t)]-\delta\omega(t)\int_t^v g(y-t)\,\rho(y-t)\,dy\right\}dt+\delta\sqrt{\omega(t)}dW^v(t) \end{aligned}$$

The characteristic function satisfies Kolmogorov's backward equation. Under the v-forward measure this equation is:

$$\begin{aligned} 0 &= \frac{\partial\psi}{\partial t}+\frac{\partial\psi}{\partial\ln P(t,T)}\left[r(t)-\frac{1}{2}\omega(t)\,A(T,T,t)+\omega(t)\,A(v,T,t)\right] \\ &\quad +\frac{1}{2}\frac{\partial^2\psi}{\partial[\ln P(t,T)]^2}\omega(t)\,A(T,T,t)+\frac{\partial\psi}{\partial\omega(t)}\{\beta[\gamma-\omega(t)]-\delta\omega(t)\,C(v,t)\} \\ &\quad +\frac{\partial\psi}{\partial r(t)}\left[\left.\frac{\partial f(t,s)}{\partial s}\right|_{s=t}-\omega(t)\,g(0)\,B(v,t)\right]-\frac{\partial^2\psi}{\partial\ln P(t,T)\,\partial\omega(t)}\omega(t)\,\delta C(T,t) \\ &\quad +\frac{1}{2}\frac{\partial^2\psi}{\partial\omega^2(t)}\delta^2\omega(t)+\frac{1}{2}\frac{\partial^2\psi}{\partial r^2(t)}\omega(t)\,g^2(0)-\frac{\partial^2\psi}{\partial\ln P(t,T)\,\partial r(t)}\omega(t)\,g(0)\,B(T,t) \\ &\quad +\frac{\partial^2\psi}{\partial r(t)\,\partial\omega(t)}\delta g(0)\,\rho(0)\,\omega(t) \end{aligned} \tag{3.16}$$

where

$$A(w,T,t) = \int_t^T\int_t^w g(u-t)\,g(y-t)\,c(u-y)\,dudy \tag{3.17}$$

$$B(w,t) = \int_t^w g(y-t)\,c(t-y)\,dy \tag{3.18}$$

$$C(w,t) = \int_t^w g(y-t)\,\rho(y-t)\,dy \tag{3.19}$$

The characteristic function solution has the following form:

$$\psi(v,t;\varphi)=e^{D(t)+J(t)r(t)+M(t)\omega(t)+i\varphi\ln P(t,T)}$$

subject to the boundary condition

$$\psi(v, m; \varphi) = e^{i\varphi \ln P(m,T)}$$

Substituting this solution into equation (3.16),

$$\begin{aligned}
0 \;=\; & \frac{dD}{dt} + \frac{dJ}{dt} r(t) + \frac{dM}{dt} \omega(t) + i\varphi \left[r(t) - \frac{1}{2} \omega(t) A(T,T,t) + \omega(t) A(v,T,t) \right] \\
& - \frac{\varphi^2}{2} \omega(t) A(T,T,t) + J(t) \left[\left. \frac{\partial f(t,s)}{\partial s} \right|_{s=t} - \omega(t) g(0) B(v,t) \right] \\
& + M(t) \{ \beta [\gamma - \omega(t)] - \delta \omega(t) C(v,t) \} + \frac{1}{2} M^2(t) \delta^2 \omega(t) + \frac{1}{2} J^2(t) \omega(t) g^2(0) \\
& - i\varphi J(t) \omega(t) g(0) B(T,t) + J(t) M(t) \delta g(0) \rho(0) \omega(t) - i\varphi M(t) \omega(t) \delta C(T,t)
\end{aligned}$$

The above equation reduces to the following system of ordinary differential equations:

$$\begin{aligned}
\frac{dM}{dt} \;=\; & i\varphi \left[\frac{1}{2} A(T,T,t) - A(v,T,t) \right] + \frac{\varphi^2}{2} A(T,T,t) + J(t) g(0) B(v,t) \\
& + M(t) \{ \beta + \delta C(v,t) \} - \frac{1}{2} M^2(t) \delta^2 - \frac{1}{2} J^2(t) g^2(0) \\
& + i\varphi J(t) g(0) B(T,t) - J(t) M(t) \delta g(0) \rho(0) + i\varphi M(t) \delta C(T,t) \\
\frac{dJ}{dt} \;=\; & -i\varphi \qquad \frac{dD}{dt} = -M(t) \beta\gamma - J(t) \left. \frac{\partial f(t,s)}{\partial s} \right|_{s=t}
\end{aligned}$$

subject to the boundary conditions $J(m) = 0, D(m) = 0, M(m) = 0$. Given a particular specification for the functions $g(s)$, $\rho(s)$, and $c(s)$, such as the following, for example,

$$g(s) = e^{-\zeta s} \qquad \rho(s) = \mu e^{-\vartheta s} \qquad c(s) = e^{-\kappa |s|} \tag{3.20}$$

the above differential equations can be solved to obtain a closed-form solution for the price of a bond option.[6]

3.2.3 Interest rate caplets

An interest rate caplet with strike rate K for the period $[s, s+\Delta]$ is a European call option on the continuously compounded forward rate $f^{\Delta}(t,s)$, which is defined as:

$$f^{\Delta}(t,s) = \frac{1}{\Delta} \int_s^{s+\Delta} f(t,u)\, du \qquad t \leq s \tag{3.21}$$

Hence,

$$d_t f^{\Delta}(t,s) = \frac{1}{\Delta} \int_s^{s+\Delta} d_t f(t,u)\, du$$

[6]The second term in the differential equation for $D(t)$ can be integrated numerically.

The option is exercised at time s if $f^{\Delta}(s,s) > K$, yielding the payoff $e^{\Delta f^{\Delta}(s,s)} - e^{\Delta K}$ at time $s+\Delta$. The price of the caplet at time t, $C_{cl}(t)$, is equal to:

$$\begin{aligned}
C_{cl}(t) &= P(t,s+\Delta)\,E_t^{\mathbb{Q}^{s+\Delta}}\left\{I_{f^{\Delta}(s,s)\geq K}\left[e^{\Delta f^{\Delta}(s,s)} - e^{\Delta K}\right]\right\} \\
&= P(t,s+\Delta)\,E_t^{\mathbb{Q}^{s+\Delta}}\left[e^{\Delta f^{\Delta}(s,s)} I_{f^{\Delta}(s,s)\geq K}\right] \\
&\quad -P(t,s+\Delta)\,e^{\Delta K} P_t^{\mathbb{Q}^{s+\Delta}}\left[f^{\Delta}(s,s)\geq K\right] \\
&= P(t,s+\Delta)\left\{\frac{\psi^{\Delta}(t,s;-i\Delta)}{2} + \frac{1}{\pi}\int_0^{\infty}\mathrm{Re}\left[\frac{e^{-i\varphi K}\psi^{\Delta}(t,s;\varphi-i\Delta)}{i\varphi}\right]d\varphi\right\} \\
&\quad -P(t,s+\Delta)\,e^{\Delta K}\left\{\frac{1}{2} + \frac{1}{\pi}\int_0^{\infty}\mathrm{Re}\left[\frac{e^{-i\varphi K}\psi^{\Delta}(t,s;\varphi)}{i\varphi}\right]d\varphi\right\}
\end{aligned}$$

where the expectation is calculated using lemma 2.3, and $\psi^{\Delta}(t,s;\varphi)$ denotes the characteristic function of $f^{\Delta}(s,s)$ under the forward measure $\mathbb{Q}^{s+\Delta}$. The dynamics of $f^{\Delta}(t,s)$ and $\omega(t)$ under $\mathbb{Q}^{s+\Delta}$ are:

$$\begin{aligned}
d_t f^{\Delta}(t,s) &= \frac{1}{\Delta}\int_s^{s+\Delta}\left[-\omega(t)\,g(u-t)\int_u^{s+\Delta} g(y-t)\,c(u-y)\,dydt\right. \\
&\quad \left. +\sqrt{\omega(t)}g(u-t)\,d_t Z^{s+\Delta}(t,u)\right]du \\
d\omega(t) &= \left\{\beta\left[\gamma-\omega(t)\right] - \delta\omega(t)\int_t^{s+\Delta} g(y-t)\,\rho(y-t)\,dy\right\}dt \\
&\quad +\delta\sqrt{\omega(t)}dW^{s+\Delta}(t)
\end{aligned}$$

The characteristic function of $f^{\Delta}(s,s)$ satisfies Kolmogorov's backward equation:

$$\begin{aligned}
0 &= \frac{\partial\psi^{\Delta}}{\partial t} - \frac{\omega(t)}{\Delta}\frac{\partial\psi^{\Delta}}{\partial f^{\Delta}(t,s)}\left[A(s+\Delta,s+\Delta,s) - A(s+\Delta,u,s)\right] \\
&\quad +\frac{\omega(t)}{2\Delta^2}\frac{\partial^2\psi^{\Delta}}{\partial\left[f^{\Delta}(t,s)\right]^2}A(s+\Delta,s+\Delta,s) + \frac{1}{\Delta}\frac{\partial^2\psi^{\Delta}}{\partial f^{\Delta}(t,s)\,\partial\omega(t)}C(s+\Delta,s)\,\delta\omega(t) \\
&\quad +\frac{\partial\psi^{\Delta}}{\partial\omega(t)}\left\{\beta\left[\gamma-\omega(t)\right] - \delta\omega(t)\,C(s+\Delta,t)\right\} + \frac{1}{2}\frac{\partial^2\psi^{\Delta}}{\partial\omega^2(t)}\omega(t)\,\delta^2
\end{aligned} \tag{3.22}$$

subject to the boundary condition

$$\psi^{\Delta}(s,s;\varphi) = e^{i\varphi f^{\Delta}(s,s)}$$

The characteristic function solution has the following form:

$$\psi^{\Delta}(t,s;\varphi) = e^{D(t)+M(t)\omega(t)+i\varphi f^{\Delta}(t,s)}$$

Substituting this solution into equation (3.22) we obtain:

$$\begin{aligned}
0 &= \frac{dD}{dt} + \frac{dM}{dt}\omega(t) - i\varphi\frac{\omega(t)}{\Delta}\left[A(s+\Delta,s+\Delta,s) - A(s+\Delta,u,s)\right] \\
&\quad -\varphi^2\frac{\omega(t)}{2\Delta^2}A(s+\Delta,s+\Delta,s) + i\varphi M(t)\,\frac{1}{\Delta}C(s+\Delta,s)\,\delta\omega(t) \\
&\quad +M(t)\left\{\beta\left[\gamma-\omega(t)\right] - \delta\omega(t)\,C(s+\Delta,t)\right\} + M^2(t)\,\frac{1}{2}\omega(t)\,\delta^2
\end{aligned}$$

which reduces to the following system of ordinary differential equations:

$$\begin{aligned} \frac{dD}{dt} &= -M(t)\,\beta\gamma \\ \frac{dM}{dt} &= U(s) + M(t)\left[W(s) + \delta C(s+\Delta,t)\right] - M^2(t)\,\frac{\delta^2}{2} \end{aligned}$$

where

$$\begin{aligned} U(s) &= i\frac{\varphi}{\Delta}\left[A(s+\Delta,s+\Delta,s) - A(s+\Delta,u,s)\right] + \frac{\varphi^2}{2\Delta^2}A(s+\Delta,s+\Delta,s) \\ W(s) &= -i\varphi\frac{1}{\Delta}C(s+\Delta,s)\,\delta + \beta \end{aligned}$$

subject to the boundary conditions $D(s) = 0$ and $M(s) = 0$. Given a particular specification for the functions $g(s)$, $\rho(s)$, and $c(s)$, the above equations can be solved to obtain a closed-form solution for the price of a caplet.

3.2.4 Interest rate spread options

A European call interest rate spread option is a security with the following payoff at maturity T:[7]

$$C(k,l,T) = \max\left[e^{r_k(T,T) - r_l(T,T)} - K, 0\right] \tag{3.23}$$

where $r_i(t,T)$ is defined as

$$r_i(t,T) = \frac{1}{i}\int_T^{T+i} f(t,u)\,du \tag{3.24}$$

In order to price interest rate spread options it is necessary to use a multifactor model in which interest rates of different maturities can be imperfectly correlated. Let $R_{kl}(t) = r_k(t,T) - r_l(t,T)$. Under the T-forward measure, the price of a spread option discounted by the price of the T-maturity discount bond is a martingale:

$$\begin{aligned} C(k,l,t) &= P(t,T)\,E_t^{Q^T}\left\{\left[e^{R_{kl}(T)} - K\right] I_{R_{kl}(T)\geq \ln K}\right\} \\ &= P(t,T)\,E_t^{Q^T}\left\{e^{R_{kl}(T)} I_{R_{kl}(T)\geq \ln K}\right\} \\ &\quad -KP(t,T)\,P_t^{Q^T}\left[R_{kl}(T) \geq \ln K\right] \\ &= P(t,T)\left\{\frac{\psi^R(t,T;-i)}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\varphi\ln K}\psi^R(t,T;\varphi - i)}{i\varphi}\right]d\varphi\right\} \\ &\quad -KP(t,T)\left\{\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\varphi\ln K}\psi^R(t,T;\varphi)}{i\varphi}\right]d\varphi\right\} \end{aligned}$$

[7]The following derivation can also be applied to the more general case of an option on a basket of yields.

where $\psi^R(t,T;\varphi)$ is the characteristic function of $R_{kl}(T)$ under the T-forward measure. The dynamics of $R_{kl}(t)$ under $\mathbb{Q}^T$ are:

$$\begin{aligned} dR_{kl}(t) &= \frac{\omega(t)}{k}\int_T^{T+k} g(u-t)\left[\int_T^u g(y-t)\,c(u-y)\,dy\right] dudt \\ &\quad -\frac{\omega(t)}{l}\int_T^{T+l} g(u-t)\left[\int_T^u g(y-t)\,c(u-y)\,dy\right] dudt \\ &\quad +\sqrt{\omega(t)}\frac{1}{k}\int_T^{T+k} g(u-t)\,d_tZ^T(t,u)\,du \\ &\quad -\sqrt{\omega(t)}\frac{1}{l}\int_T^{T+l} g(u-t)\,d_tZ^T(t,u)\,du \end{aligned}$$

Kolmogorov's backward equation for the characteristic function of $R_{kl}(T)$ is:

$$\begin{aligned} 0 &= \frac{\partial\psi^R}{\partial t}+\frac{\partial\psi^R}{\partial R_{kl}(t)}\omega(t)\left[\frac{1}{k}A(T+k,u,T)-\frac{1}{l}A(T+l,u,T)\right] \\ &\quad +\frac{1}{2}\frac{\partial^2\psi^R}{\partial R_{kl}^2(t)}\omega(t)\left[\frac{1}{k^2}A(T+k,T+k,T)+\frac{1}{l^2}A(T+l,T+l,T)\right. \\ &\quad \left.-\frac{2}{kl}A(T+k,T+l,T)\right]+\frac{\partial\psi^R}{\partial\omega(t)}\{\beta[\gamma-\omega(t)]-\delta\omega(t)\,C(T,t)\} \\ &\quad +\frac{\delta^2}{2}\omega(t)\frac{\partial^2\psi^R}{\partial\omega^2(t)}+\frac{\partial^2\psi^R}{\partial R_{kl}(t)\,\partial\omega(t)}\omega(t)\,\delta\left[\frac{1}{k}C(T+k,T)-\frac{1}{l}C(T+l,T)\right] \end{aligned}$$

The characteristic function solution has the following form:

$$\psi^R(t,T;\varphi)=e^{D(t)+M(t)\omega(t)+i\varphi R_{kl}(t)}$$

subject to the boundary condition:

$$\psi^R(T,T;\varphi)=e^{i\varphi R_{kl}(T)}$$

The functions $D(t)$ and $M(t)$ are the solutions to the following system of ordinary differential equations:

$$\begin{aligned} \frac{dD}{dt} &= -M(t)\,\beta\gamma \\ \frac{dM}{dt} &= -i\varphi\left[\frac{1}{k}A(T+k,u,T)-\frac{1}{l}A(T+l,u,T)\right]+M(t)\,[\beta+\delta C(T,t)] \\ &\quad +\frac{\varphi^2}{2}\left[\frac{1}{k^2}A(T+k,T+k,T)+\frac{1}{l^2}A(T+l,T+l,T)-\frac{2}{kl}A(T+k,T+l,T)\right] \\ &\quad -\frac{\delta^2}{2}M^2(t)-i\varphi M(t)\,\delta\left[\frac{1}{k}C(T+k,T)-\frac{1}{l}C(T+l,T)\right] \end{aligned}$$

subject to the boundary conditions $D(T)=0$ and $M(T)=0$.

3.2.5 The model under the physical measure

The assumption of stochastic forward rate volatility implies that markets are incomplete.[8] As a result, bonds are not enough to span all interest rate contingent claims, and there exist an infinite number of equivalent martingale measures (EMMs), each corresponding to a different specification for the volatility risk premium.

We assume that the dynamics of forward rates under the physical measure $\mathbb{P}$ are as follows:

$$d_t f(t,s) = \eta(t,s)\,dt + \sigma(t,s)\,d_t Z^{\mathbb{P}}(t,s) \tag{3.25}$$

where $Z^{\mathbb{P}}(t,s)$ is a random field satisfying assumption (I) under the physical measure.

The dynamics of bond prices under the physical measure are:

$$\begin{aligned}\frac{d_t P(t,s)}{P(t,s)} &= r(t)\,dt - \int_t^s \eta(t,y)\,dy dt - \int_t^s \sigma(t,y)\,d_t Z^{\mathbb{P}}(t,y)\,dy \\ &+ \frac{1}{2}\int_t^s\int_t^s \sigma(t,y)\,\sigma(t,x)\,c(x-y)\,dx dy dt \end{aligned} \tag{3.26}$$

We assume the existence of a stochastic discount factor $M(t)$ with dynamics:

$$\frac{dM(t)}{M(t)} = -r(t)\,dt - \int_t^\infty \lambda(t,y)\,d_t Z^{\mathbb{P}}(t,y)\,dy - \psi(t)\,dW^{\mathbb{P}}(t) \tag{3.27}$$

where $\lambda(t,s)$ and $\psi(t)$ denote the forward rate and volatility risk premia respectively. Let $P_M(t,s) = M(t)P(t,s)$ denote the time-t deflated price of a pure discount bond with maturity s. The dynamics of $P_M(t,s)$ are:

$$\begin{aligned}\frac{dP_M(t,s)}{P_M(t,s)} &= \left\{\frac{1}{2}\int_t^s\int_t^s \sigma(t,y)\,\sigma(t,x)\,c(x-y)\,dx dy - \int_t^s \eta(t,y)\,dy\right\} dt \\ &+ \int_t^s \left[\int_t^\infty \lambda(t,u)\,c(u-y)\,du\right] \sigma(t,y)\,dy dt \\ &+ \psi(t)\int_t^s \sigma(t,y)\,\rho(y-t)\,dy dt - \int_t^\infty \lambda(t,y)\,d_t Z^{\mathbb{P}}(t,y)\,dy \\ &- \psi(t)\,dW^{\mathbb{P}}(t) - \int_t^s \sigma(t,y)\,d_t Z^{\mathbb{P}}(t,y)\,dy\end{aligned}$$

Under the assumption of no arbitrage opportunities, $P_M(t,s)$ must be a martingale. Hence,

$$\begin{aligned}\int_t^s \eta(t,y)\,dy &= \frac{1}{2}\int_t^s\int_t^s \sigma(t,y)\,\sigma(t,x)\,c(x-y)\,dx dy + \psi(t)\int_t^s \sigma(t,y)\,\rho(y-t)\,dy \\ &+ \int_t^s \left[\int_t^\infty \lambda(t,u)\,c(u-y)\,du\right] \sigma(t,y)\,dy\end{aligned}$$

[8] Similarly to the previous chapter, markets are incomplete when forward rate volatility is stochastic, unless one of the following conditions holds: (a) forward rate volatility risk can be traded, or (b) aggregate consumption is uncorrelated with forward rate volatility, so that the risk premium associated with volatility under the physical measure is zero.

Differentiating with respect to the maturity date s,

$$\begin{aligned}\eta(t,s) &= \sigma(t,s)\int_t^s \sigma(t,x)\,c(x-s)\,dx + \sigma(t,s)\int_t^\infty \lambda(t,u)\,c(u-s)\,du \\ &\quad +\psi(t)\,\sigma(t,s)\,\rho(s-t)\end{aligned}$$

The change of measure from the physical to the risk-neutral is therefore given by:

$$d_t Z(t,s) = \left[\int_t^\infty \lambda(t,u)\,c(u-s)\,du + \psi(t)\,\rho(s-t)\right]dt + d_t Z^{\mathbb{P}}(t,s) \tag{3.28}$$

We define the process $\tilde{m}(t)$ as follows:

$$\frac{d\tilde{m}(t)}{\tilde{m}(t)} = \int_t^\infty \lambda(t,y)\,d_t Z(t,y)\,dy + \psi(t)\,dW(t) \tag{3.29}$$

Under the physical measure, $W^{\mathbb{P}}(t)$ defined below is a standard Brownian motion and satisfies assumption (II) jointly with $Z^{\mathbb{P}}(t,s)$:

$$\begin{aligned}dW^{\mathbb{P}}(t) &= dW(t) - \frac{1}{\tilde{m}(t)}d\langle W(t),\tilde{m}(t)\rangle \\ &= dW(t) - \left[\int_t^\infty \lambda(t,y)\,\rho(y-t)\,dy + \psi(t)\right]dt \end{aligned} \tag{3.30}$$

Hence, the dynamics of $\omega(t)$ under the physical measure are:

$$\begin{aligned}d\omega(t) &= \left\{\beta\left[\gamma-\omega(t)\right] + \delta\sqrt{\omega(t)}\left[\int_t^\infty \lambda(t,y)\,\rho(y-t)\,dy + \psi(t)\right]\right\}dt \\ &\quad +\delta\sqrt{\omega(t)}dW^{\mathbb{P}}(t)\end{aligned} \tag{3.31}$$

3.3 Random-field models for LIBOR and swap rates with stochastic volatility

3.3.1 Introduction

An important drawback of the HJM approach from a practical viewpoint is that the interest rate being modelled is unobservable. In response to this problem market models of the term structure emerged. This class of models applies the HJM approach to observable and discretely compounded interest rates. Initial market models of forward LIBOR (Brace, Gatarek, and Musiela, 1997; Musiela and Rutkowski, 1997; Miltersen, Sandmann, and Sondermann, 1997) and swap rates (Jamshidian, 1997) assume that LIBOR and swap rates, respectively, follow a lognormal process under the corresponding forward measure. These LIBOR and swap rate models are in fact mutually inconsistent, because these two types of interest rates cannot be simultaneously lognormally distributed under the no-arbitrage condition.

Under the lognormality assumption, the prices for caplets and swaptions are given by a formula which is very similar to the Black-Scholes formula for stock options and is known as the Black formula. This formula was originally derived by Black (1976) as the price of an option on commodity futures when the change in the futures price is assumed to be lognormally distributed. The use of the Black formula by traders for pricing interest rate derivatives was at first considered erroneous, as forward rates for different maturities cannot all be simultaneously lognormal under the same measure without arbitrage opportunities arising. However, it was subsequently shown by Brace, Gatarek, and Musiela (1997) that market practice is consistent with theory if each forward rate is assumed to be lognormal under the corresponding forward measure. Similarly to the Black-Scholes formula, the Black formula can be inverted to obtain the implied volatility of caplets and swaptions.

The existence of implied volatility smiles and skews in interest rate derivatives markets has prompted the appearance of market models with stochastic volatility. Andersen and Brotherton-Ratcliffe (2001), Joshi and Rebonato (2001), and Andersen and Andreasen (2002) extend the standard LIBOR market model to include stochastic volatility. Jarrow, Li, and Zhao (2003) compare empirically the standard lognormal LIBOR market model against the stochastic volatility model of Andersen and Brotherton-Ratcliffe (2001), the market model with jumps of Glasserman and Kou (2003),[9] and a combination of the two. They find that while the models with stochastic volatility and/or jumps perform better than the standard LIBOR market model, none of the models considered can adequately explain the entire cap implied volatility smile. In this section we construct random field market models for forward LIBOR and swap rates with stochastic volatility.

3.3.2 LIBOR model

The forward LIBOR rate $L(t,T_j)$ is defined as the interest rate that is earned on a risk-free loan between dates T_j and T_{j+1}:

$$1 + a_{T_j} L(t,T_j) = e^{\int_{T_j}^{T_{j+1}} f(t,x)dx} \qquad t \leq T_j < T_{j+1} \tag{3.32}$$

where a_{T_j} is the *daycount fraction* for the time period $[T_j, T_{j+1}]$. Hence,

$$L(t,T_j) = \frac{1}{a_{T_j}} \left[\frac{P(t,T_j)}{P(t,T_{j+1})} - 1 \right] \tag{3.33}$$

[9]In their model volatility is deterministic but there are jumps in interest rates.

Let $U(t) = \int_{T_j}^{T_{j+1}} f(t,x)\,dx$. Then $L(t,T_j) = \frac{1}{a_{T_j}}\left[e^{U(t)} - 1\right]$ and

$$\begin{aligned} dU(t) &= \int_{T_j}^{T_{j+1}} d_t f(t,x)\,dx \\ d_t L(t,T_j) &= \frac{1 + a_{T_j} L(t,T_j)}{a_{T_j}} \left[dU(t) + \frac{1}{2} d\langle U(t)\rangle \right] \\ dU(t) + \frac{1}{2} d\langle U(t)\rangle &= \int_{T_j}^{T_{j+1}} \sigma(t,x)\, d_t Z(t,x)\,dx \\ &\quad + \int_{T_j}^{T_{j+1}} \sigma(t,x) \int_t^x \sigma(t,y)\, c(x-y)\,dydxdt \\ &\quad + \frac{1}{2} \int_{T_j}^{T_{j+1}} \int_{T_j}^{T_{j+1}} \sigma(t,x)\,\sigma(t,y)\, c(x-y)\,dydxdt \end{aligned}$$

Let

$$\gamma_j(t) = \frac{1 + a_{T_j} L(t,T_j)}{a_{T_j} L(t,T_j)} \tag{3.34}$$

The dynamics of $L(t,T_j)$ under the forward measure $\mathbb{Q}^{T_{j+1}}$ are:

$$d_t L(t,T_j) = L(t,T_j)\,\gamma_j(t) \int_{T_j}^{T_{j+1}} \sigma(t,x)\, d_t Z^{T_{j+1}}(t,x)\,dx \tag{3.35}$$

Hence $L(t,T_j)$ is a martingale under $\mathbb{Q}^{T_{j+1}}$. The standard LIBOR market model assumes that forward LIBOR rates have the following dynamics under the corresponding forward measure:

$$d_t L(t,T_j) = \sigma(t,T_j)\, L(t,T_j)\, dW^{T_{j+1}}(t)$$

where $\sigma(t,T_j)$ is a deterministic function of t and T_j, and $W^{T_{j+1}}(t)$ is a standard Brownian motion under $\mathbb{Q}^{T_{j+1}}$. Andersen and Brotherton-Ratcliffe (2001), and Andersen and Andreasen (2002) extend the standard LIBOR market model to include stochastic volatility. They assume that $\sigma(t,T_j)$ has a stochastic component $V(t)$ which is modelled as follows:

$$dV(t) = \kappa\,[\theta - V(t)]\,dt + \eta f\,[V(t)]\, dW_V^{T_{j+1}}(t)$$

where $W_V^{T_{j+1}}(t)$ is a standard Brownian motion under $\mathbb{Q}^{T_{j+1}}$ which is independent of the Brownian motion driving LIBOR rates. Andersen and Brotherton-Ratcliffe (2001) derive the price of a caplet using the Taylor series expansion technique of Hull and White (1987). Andersen and Andreasen (2002) price caplets and swaptions using Heston's (1993) inverse Fourier transform method.

In this section we assume that the dynamics of $L(t,T_j)$ under the forward measure $\mathbb{Q}^{T_{j+1}}$ are:[10]

$$d_t L(t,T_j) = L(t,T_j)\,\sqrt{\varpi(t)}\, g(T_j - t)\, d_t Z^{T_{j+1}}(t,T_j) \tag{3.36}$$

[10]The random field $Z(t,s)$ here is not related to the one used earlier.

where $g(u)$ is a deterministic function of u, and $\varpi(t)$ is a stochastic process with the following dynamics under $\mathbb{Q}^{T_{j+1}}$:

$$d\varpi(t) = \beta[\gamma - \varpi(t)]\,dt + \delta\sqrt{\varpi(t)}dW^{T_{j+1}}(t) \tag{3.37}$$

where $W^{T_{j+1}}(t)$ is a standard Brownian motion under $\mathbb{Q}^{T_{j+1}}$, and $Z^{T_{j+1}}(t, T_j)$, $W^{T_{j+1}}(t)$ satisfy assumptions (I) and (II) under $\mathbb{Q}^{T_{j+1}}$. This specification of the dynamics of LIBOR rates is not compatible with the forward rate model described earlier, but it allows us to obtain a closed-form solution for the price of a LIBOR caplet.

A LIBOR caplet with strike rate K for the period $[T_j, T_{j+1}]$ is a European call option on the LIBOR rate $L(t, T_j)$. The payoff of the caplet at time T_{j+1} is equal to:

$$C_{cl}(T_{j+1}) = a_{T_j}\max[L(T_j, T_j) - K, 0]$$

Under the forward measure $\mathbb{Q}^{T_{j+1}}$ the price of the caplet at time t is:

$$\begin{aligned}
C_{cl}(t) &= P(t, T_{j+1})\,a_{T_j}E_t^{\mathbb{Q}^{T_{j+1}}}\left\{I_{L(T_j,T_j)\geq K}[L(T_j, T_j) - K]\right\} \\
&= P(t, T_{j+1})\,a_{T_j}E_t^{\mathbb{Q}^{T_{j+1}}}\left[L(T_j, T_j)\,I_{L(T_j,T_j)\geq K}\right] \\
&\quad -P(t, T_{j+1})\,a_{T_j}KP_t^{\mathbb{Q}^{T_{j+1}}}[L(T_j, T_j) \geq K] \\
&= P(t, T_{j+1})\,a_{T_j}\left\{\frac{\psi^L(t, T_j; -i)}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\varphi\ln K}\psi^L(t, T_j; \varphi - i)}{i\varphi}\right]d\varphi\right\} \\
&\quad -P(t, T_{j+1})\,a_{T_j}K\left\{\frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\varphi\ln K}\psi^L(t, T_j; \varphi)}{i\varphi}\right]d\varphi\right\}
\end{aligned}$$

where $\psi^L(t, T_j; \varphi)$ denotes the characteristic function of $\ln L(T_j, T_j)$ under measure $\mathbb{Q}^{T_{j+1}}$. The dynamics of $\ln L(t, T_j)$ under $\mathbb{Q}^{T_{j+1}}$ are:

$$d_t \ln L(t, T_j) = \sqrt{\varpi(t)}g(T_j - t)\,d_tZ^{T_{j+1}}(t, T_j) - \frac{1}{2}\varpi(t)\,g^2(T_j - t)\,dt \tag{3.38}$$

Kolmogorov's backward equation for the characteristic function is:

$$\begin{aligned}
0 &= \frac{\partial\psi^L}{\partial t} + \frac{\varpi(t)}{2}g^2(T_j - t)\left[\frac{\partial^2\psi^L}{\partial[\ln L(t, T_j)]^2} - \frac{\partial\psi^L}{\partial\ln L(t, T_j)}\right] + \frac{\partial\psi^L}{\partial\varpi(t)}\beta[\gamma - \varpi(t)] \\
&\quad +\frac{\delta^2}{2}\frac{\partial^2\psi^L}{\partial\varpi^2(t)}\varpi(t) + \frac{\partial^2\psi^L}{\partial\ln L(t, T_j)\,\partial\varpi(t)}\delta\varpi(t)\,g(T_j - t)\,\rho(T_j - t)
\end{aligned} \tag{3.39}$$

The characteristic function solution has the following form:

$$\psi^L(t, T_j; \varphi) = e^{D(t) + M(t)\varpi(t) + i\varphi\ln L(t, T_j)}$$

subject to the boundary condition

$$\psi^L(T_j, T_j; \varphi) = e^{i\varphi\ln L(T_j, T_j)}$$

Substituting this solution into equation (3.39), we obtain the following system of ordinary differential equations, subject to the boundary conditions $D(T_j) = 0$ and $M(T_j) = 0$,

$$\begin{aligned}\frac{dD}{dt} &= -M(t)\beta\gamma \\ \frac{dM}{dt} &= \frac{\varphi}{2}g^2(T_j - t)(i+\varphi) + M(t)\beta - \frac{\delta^2}{2}M^2(t) - i\varphi M(t)\delta g(T_j - t)\rho(T_j - t)\end{aligned}$$

which yield a closed-form solution for the price of a caplet, given a particular specification for the functions $g(u)$ and $\rho(u)$.

3.3.3 Swap rate model

An interest rate swap is a contract where two parties agree to exchange interest rate payments. One party pays fixed interest while the other pays floating interest. A floating interest payment made at time T_{i+1} is based on the LIBOR rate fixed at time T_i, and is equal to $a_{T_i}L(T_i, T_i)$. The par swap rate F is defined as the fixed interest rate for which the present value of the swap is equal to zero. Let T_n denote the date at which the first LIBOR rate is fixed, and $T_{n+1}, \ldots, T_N$ the dates at which swap payments are exchanged. The forward par swap rate is given by:

$$F(t, T_n, T_N) = \frac{P(t, T_n) - P(t, T_N)}{\sum_{i=n+1}^{N} a_{i-1}P(t, T_i)} \tag{3.40}$$

Let $\mathbb{Q}^{n+1,N}$ denote the measure under which asset prices discounted by the following numeraire are martingales:

$$D(t, T_{n+1}, T_N) = \sum_{i=n+1}^{N} a_{i-1}P(t, T_i) \tag{3.41}$$

$\mathbb{Q}^{n+1,N}$ is the forward swap measure (Jamshidian, 1997). The par swap rate $F(t, T_n, T_N)$ is also a martingale under $\mathbb{Q}^{n+1,N}$.

A swaption gives the right to enter into a swap with fixed rate K at expiration. Let $S(t, T_n, T_N)$ denote the time-t price of a payer swaption, i.e. a European option on a swap in which the buyer of the option pays fixed interest if the option is exercised. The option expires at time T_n, and if it is exercised, the swap starts at time T_n and ends at T_N. The present value of the swap is:

$$V(t, T_n, T_N) = P(t, T_n) - P(t, T_N) - K\sum_{i=n+1}^{N} a_{i-1}P(t, T_i)$$

Thus a swaption is an option on a portfolio of bonds. The price of the swaption at time t is:

$$\begin{aligned}
S(t,T_n,T_N) &= D(t,T_{n+1},T_N)\,E_t^{\mathbb{Q}^{n+1,N}}\left\{\frac{V(T_n,T_n,T_N)}{D(T_n,T_{n+1},T_N)}I_{F(T_n,T_n,T_N)\geq K}\right\}\\
&= D(t,T_{n+1},T_N)\,E_t^{\mathbb{Q}^{n+1,N}}\left\{[F(T_n,T_n,T_N)-K]\,I_{F(T_n,T_n,T_N)\geq K}\right\}\\
&= D(t,T_{n+1},T_N)\,E_t^{\mathbb{Q}^{n+1,N}}\left[F(T_n,T_n,T_N)\,I_{F(T_n,T_n,T_N)\geq K}\right]\\
&\quad -KD(t,T_{n+1},T_N)\,P_t^{\mathbb{Q}^{n+1,N}}\left[F(T_n,T_n,T_N)\geq K\right]\\
&= D(t,T_{n+1},T_N)\left\{\frac{\psi^F(t,T_n,T_N;-i)}{2}\right.\\
&\quad\left.+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\varphi\ln K}\psi^F(t,T_n,T_N;\varphi-i)}{i\varphi}\right]d\varphi\right\}\\
&\quad -KD(t,T_{n+1},T_N)\left\{\frac{1}{2}+\frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[\frac{e^{-i\varphi\ln K}\psi^F(t,T_n,T_N;\varphi)}{i\varphi}\right]d\varphi\right\}
\end{aligned}$$

where $\psi^F(t,T_n,T_N;\varphi)$ denotes the characteristic function of $\ln F(T_n,T_n,T_N)$ under $\mathbb{Q}^{n+1,N}$. We assume that the dynamics of the forward par swap rate under $\mathbb{Q}^{n+1,N}$ are as follows:

$$d_tF(t,T_n,T_N) = F(t,T_n,T_N)\sqrt{v(t)}g(T_n-t,T_N-t)\,d_tZ^{n+1,N}(t,T_n,T_N) \tag{3.42}$$

where $g(x,y)$ is a deterministic function of x and y, and $v(t)$ is a stochastic process with dynamics

$$dv(t) = \beta[\gamma-v(t)]\,dt+\delta\sqrt{v(t)}dW^{n+1,N}(t) \tag{3.43}$$

where $W^{n+1,N}(t)$ is a standard Brownian motion under $\mathbb{Q}^{n+1,N}$, and $Z^{n+1,N}(t,T_n,T_N)$, $W^{n+1,N}(t)$ satisfy assumptions (I) and (II) under $\mathbb{Q}^{n+1,N}$.

Kolmogorov's backward equation for $\psi^F(t,T_n,T_N;\varphi)$ and its solution are similar to those derived in the previous section for LIBOR caplets:

$$\psi^F(t,T_n,T_N;\varphi) = e^{D(t)+M(t)v(t)+i\varphi\ln F(t,T_n,T_N)}$$

where $D(t)$ and $M(t)$ are the solutions to the following system of ordinary differential equations:

$$\begin{aligned}
\frac{dD}{dt} &= -M(t)\,\beta\gamma\\
\frac{dM}{dt} &= \frac{\varphi}{2}g^2(T_n-t,T_N-t)(i+\varphi)+M(t)\,\beta-\frac{\delta^2}{2}M^2(t)\\
&\quad -i\varphi M(t)\,\delta g(T_n-t,T_N-t)\,\rho(T_n-t,T_N-t)
\end{aligned}$$

subject to the boundary conditions $D(T_n)=0$ and $M(T_n)=0$.

3.4 Conclusion

In this chapter we construct random field models of the term structure of interest rates with stochastic volatility. We assume that the volatility of interest rates can be decomposed into a deterministic function of the time to maturity and a maturity-independent stochastic process driven by a standard Brownian motion. We allow for non-zero correlation between this Brownian motion and the random field driving interest rates. The separability of the volatility function enables us to obtain closed-form solutions for the prices of bond options, interest rate caplets, interest rate spread options, and swaptions using Heston's (1993) inverse Fourier transform method.

We present three different models for instantaneous forward rates, forward LIBOR, and swap rates. These models fit initial instantaneous forward rate, forward LIBOR, and swap rate curves, respectively. While these models are mutually inconsistent, each one can be useful for valuing different types of interest rate contingent claims. Having a model with simple dynamics for the interest rate that is relevant for pricing a particular derivative has advantages in terms of ease of calibration and accuracy in pricing. These factors can be more important than having a single model for all securities, in practice, as inconsistencies between prices in different markets may exist due to market imperfections.

Chapter 4

Estimation of random-field models of the forward rate with stochastic volatility

4.1 Introduction

Empirical research on the term structure of interest rates has largely focused on models with a small number of factors, due to the difficulties encountered in estimating large-scale multifactor models. A popular multifactor term structure model is the Heath, Jarrow, and Morton (1992) (HJM) model. Pearson and Zhou (1999) estimate non-parametrically the forward rate volatility functions in a HJM term structure model with one and two factors. They find that the relationship between forward rate volatility and the level of forward rates is weak, while there is a strong relationship between volatility and the slope of the forward rate curve for low and moderate forward rate levels. Jeffrey *et al.* (2003) develop a non-parametric estimator for the volatility function of the yield curve in a HJM model which takes into account yield curve measurement errors, and they implement this estimator using HJM models with one and two factors. Chiarella, Pasquali, and Runggaldier (2001) propose a Bayesian filtering algorithm for estimating a HJM model where forward rate volatility is a function of the time to maturity, the spot rate, and one fixed maturity forward rate. In all of these three empirical papers the volatility of forward rates is assumed to be deterministic.

Random field forward rate models extend the HJM framework to an infinite number of factors. As this is a relatively recent approach of modelling the term structure of interest rates, there exist few empirical studies of such models. Santa-Clara and Sornette

(2001) calibrate a deterministic volatility random field forward rate model using two types of random field shocks: an Ornstein-Uhlenbeck random field and a subexponential correlation random field. In each case the model is described by three parameters, which are estimated by minimizing the sum of squared differences between empirical and model-implied bond price covariances for all available maturities.

Longstaff, Santa-Clara, and Schwartz (2001a) use a random field LIBOR model with deterministic volatility to examine the relative valuation of caps and swaptions. They find that long-dated swaptions are undervalued relative to other swaptions and cap prices differ from those implied by swaption prices under the no-arbitrage condition. They suggest that a model which allows for a time-varying covariance structure may explain better the observed relative pricing of caps and swaptions. In another paper Longstaff *et al.* (2001b) calculate the loss in using a single-factor model to price American swaptions when the true model is a random field model, and they find that using a misspecified model can be very costly in this case. In both of these studies Longstaff *et al.* use principal components analysis to decompose the forward LIBOR covariance matrix into a number of factors. The number of significant factors found is four and, as a result, only four factors are used to approximate the original infinite-factor random field LIBOR model by a finite-factor one. This approximation, however, reduces their random field LIBOR model to the standard LIBOR market model introduced by Brace *et al.* (1997). In fact, Kerkhof and Pelsser (2002) show that when random field LIBOR models are reduced to a finite number of factors, they become observationally equivalent to the standard LIBOR market model. By observational equivalence Kerkhof and Pelsser (2002) mean that

> "for every specification in the class of discrete string models one can find a specification in the class of market models with the same probabilistic properties and vice versa."

In this chapter we estimate random field models of the forward rate while maintaining their infinite-factor structure. Three random field term structure models are considered. In model I forward rate volatility is a function of the spot rate and the time to maturity. In model II forward rate volatility is the product of a deterministic function of the time to maturity and a stochastic process driven by a Brownian motion. In model III volatility is driven by a random field. In the two stochastic volatility models we assume that the stochastic process driving volatility is uncorrelated with the random field driving forward rates. The number of model parameters is five, seven, and ten in models I, II, and III respectively. Each model is estimated using seven years of daily UK and US forward rate data, spanning times to maturity between zero and 120 months at increments of one month. Model I can be estimated using standard estimation techniques, such as the generalized method of moments and maximum likelihood. The combination of a large number of factors with stochastic volatility in models II and III, however, requires a more careful selection of the estimation method.

A number of different approaches have been developed for estimating stochastic volatility models: applications of the generalized method of moments (GMM) introduced by Hansen (1982) (Wiggins, 1987; Scott, 1987; Melino and Turnbull, 1990; Andersen, 1994; Ho, Perraudin, and Sorensen, 1996), Monte Carlo maximum likelihood (MCML) (Danielsson and Richard, 1993), nonparametric simulated maximum likelihood (Fermanian and Salanie, 2004), the simulation-based method of moments (SMM) (Duffie and Singleton, 1993; Andersen and Sorensen, 1996), the indirect inference approach (Gourieroux, Monfort, and Renault, 1993), the efficient method of moments (EMM) (Gallant and Tauchen, 1996; Andersen and Lund, 1997; Gallant and Long, 1997), Kalman filter techniques (Harvey, Ruiz, and Shephard, 1994), and Markov chain Monte Carlo methods (MCMC) (Jacquier, Polson, and Rossi, 1994).

Applications of the above methods have concentrated on models with a low number of latent variables, due to the difficulties presented by high-dimensional stochastic volatility models. Nonparametric simulated maximum likelihood is unfeasible for multivariate models of very high dimension – 121 in our case – as the sample size required for each simulation becomes prohibitively large.[1] Method-of-moments-based techniques, such as GMM and SMM, also become computationally unfeasible as the number of latent variables and model parameters increases, and are inefficient, in general, unless they involve a continuum of moments (Carrasco and Florens, 2002) or other special cases. Recent advances in GMM estimation based on the characteristic function (Singleton, 2001; Jiang and Knight, 2002; Chacko and Viceira, 2003) require the characteristic function of the observations to be known in closed form, which in our case is not available in any of the models we consider. Carrasco *et al.* (2004) extend this approach to cases where the characteristic function is not known analytically, by estimating it using simulations. In this way, however, discretization bias is introduced, while the computational cost increases exponentially with the dimension of the model.

MCMC estimation methods offer an alternative avenue for handling complex stochastic volatility models. They also allow the estimation of stochastic volatility and the comparison of non-nested models using the Bayes factor. The success of MCMC

[1]See Silverman (1986), pp. 93–94, and Scott (1992), chapter 7.

methods, however, depends on the rate of convergence of the estimated parameters to the true values of the model parameters. As the number of latent variables increases – which in our case is equal to the number of observed variables, 121, in model III – the number of iterations required for convergence to occur can become unfeasibly large.[2]

In recent years a number of studies have addressed the problem of discretization bias in the estimation of continuous-time models using discrete-time data. Ait-Sahalia (2002) estimates fully-observed continuous-time diffusions through maximum likelihood, by approximating the true likelihood function using Hermite polynomials. Ait-Sahalia and Kimmel (2004) apply this method to the estimation of stochastic volatility stock price models, by assuming that stock price volatility can be approximated by the implied volatility of close-to-maturity at-the-money options, and using closed-form expansions for the joint likelihood function of the stock price and its volatility. Pedersen (1995), Brandt and Santa-Clara (2002), and Durham and Gallant (2002) develop a method of maximum likelihood estimation of diffusions where the true likelihood function is approximated through simulations that serve to integrate out unobserved states between observations. This technique, known as simulated maximum likelihood, can also be extended to latent variable models. A related approach is suggested by Elerian, Chib, and Shephard (2001), and Eraker (2001), who employ MCMC methods to estimate diffusions, using simulations to generate additional data between observations and integrating them out of the likelihood function. Simulation-based methods, however, increase the computation time of the estimation process, which can make the procedure impractical for high-dimensional models. Using high-frequency data may be a better way to reduce discretization bias in such cases. We use daily data in order to minimize the discretization bias.

We estimate model I using maximum likelihood,[3] and models II and III using MCML, adopting a standard Euler discretization scheme. In MCML volatility is integrated out of the likelihood function using simulations. The accuracy of the likelihood function estimate depends on the variance of the Monte Carlo estimator. There exist a number of variance reduction techniques, one of which is importance sampling. In this method volatility samples are not drawn from the volatility density implied by the model, but from the importance sampling density, which is defined in such a way that the ratio

[2]Chib *et al.* (2002) estimate a high-dimensional stochastic volatility model using MCMC methods, but their technique relies on the particular specification of the model, which allows them to speed up the sampling process through a transformation of the model into a number of low-dimensional independent models. In our case, however, such a simplifying transformation is not possible.

[3]The asymptotic efficiency of maximum likelihood estimation is well-known. See, for example, Rao (1973).

of the joint density of forward rates and volatility to the importance sampling density is almost constant for all volatility values. A careful construction of the importance sampling density can result in a substantial reduction in the variance of the Monte Carlo estimator. The rest of this chapter is organized as follows: Section 4.2 describes the random field forward rate models to be estimated. Section 4.3 explains the estimation procedure for each model. Section 4.4 describes the dataset and section 4.5 presents the empirical results. Finally, section 4.6 concludes and provides suggestions for further research.

4.2 The models

Uncertainty is modelled as a complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$, where $\mathbb{Q}$ represents the risk-neutral measure and $\mathbb{F} = \{\mathcal{F}_t\}_{t \geq 0}$ is the $\mathbb{Q}$-augmentation of the following natural filtration in each model:

$$\begin{aligned}
\text{model I} &: \mathcal{F}_t = \sigma\left(Z_f(u,s), 0 \leq u \leq t, s \geq 0\right) \\
\text{model II} &: \mathcal{F}_t = \sigma\left(Z_f(u,s), W(u), 0 \leq u \leq t, s \geq 0\right) \\
\text{model III} &: \mathcal{F}_t = \sigma\left(Z_f(u,s), Z_\sigma(u,s), 0 \leq u \leq t, s \geq 0\right)
\end{aligned}$$

where $Z_f(t,s)$ and $Z_\sigma(t,s)$ are random fields, $W(t)$ is a standard Brownian motion, t denotes current time and s the time to maturity, $t, s \geq 0$. We assume that $\mathbb{F}$ satisfies the *usual conditions.*

At each time t there exist zero-coupon risk-free bonds for all times to maturity $s \geq 0$. The time-t price of a zero-coupon risk-free bond which pays one currency unit at time $t+s$ is denoted by $P(t, t+s)$. The instantaneous forward rate f at time t for time to maturity s is defined as:

$$f(t,s) = -\frac{\partial \ln P(t, t+s)}{\partial s} \tag{4.1}$$

and the spot rate at time t is $r(t) = f(t,0)$. We assume that the dynamics of the instantaneous forward rate under the risk-neutral measure are as follows:

$$d_t f(t,s) = \alpha(t,s)\, dt + \sigma(t,s)\, d_t Z_f(t,s) \tag{4.2}$$

Table 4.1 presents the specification for the volatility function $\sigma(t,s)$ in each model. In models I and II $g(s)$ denotes a deterministic function of the time to maturity s. Model I is a random field forward rate model with deterministic volatility that belongs to the class considered by Goldstein (2000). In model II forward rate volatility is the

Model	Volatility function
I	$\sigma(t,s) = g(s)\sqrt{r(t)}$
II	$\sigma(t,s) = \sigma(t)\,g(s)$ $d\ln\sigma(t) = \zeta\left[\theta - \ln\sigma(t)\right]dt + \eta dW(t)$
III	$d_t \ln\sigma(t,s) = \zeta(s)\left[\theta(s) - \ln\sigma(t,s)\right]dt + \eta(s)\,d_t Z_\sigma(t,s)$

Table 4.1: Volatility function specification

product of a deterministic function of the time to maturity and a maturity-independent stochastic process $\sigma(t)$ driven by a standard Brownian motion. In model III volatility is driven by a random field $Z_\sigma(t,s)$, assumed to be Gaussian. In all three models the random field $Z_f(t,s)$ driving forward rates is also assumed to be Gaussian. In models II and III we assume that the random shocks in volatility are uncorrelated with those of forward rates:

ASSUMPTION I *(i)* $d_t Z_f(t,s)$ *and* $dW(t)$ *are independent.*

(ii) $d_t Z_f(t,s_1)$ *and* $d_t Z_\sigma(t,s_2)$ *are independent* $\forall\ s_1,\ s_2$.

ASSUMPTION II *The random fields* $Z_i(t,s)$, $i = f, \sigma$, *satisfy the following conditions:*

(i) $Z_i(t,s)$ *is continuous in* t *and* s.

(ii) $E\left[d_t Z_i(t,s)\right] = 0$.

(iii) $var\left[d_t Z_i(t,s)\right] = dt$.

(iv) $cor\left[d_t Z_i(t,s_1), d_t Z_i(t,s_2)\right] = C_i(s_1,s_2)$.

(v) $d_t Z_i(t_1,s_1)$ *and* $d_t Z_i(t_2,s_2)$ *are independent* $\forall\ t_1 \neq t_2$.

There exists a bank account with price process $B(t)$ adapted to the filtration $\{\mathcal{F}_t\}$, and we take $B(0)$ to be equal to one. The value of the bank account at time t is $B(t) = e^{\int_0^t r(x)dx}$. Hence, $dB(t) = B(t)\,r(t)\,dt$.

Under the no-arbitrage condition on bond prices, the forward rate drift is equal to (see appendix E):

$$\alpha(t,s) = \frac{\partial f(t,s)}{\partial s} + \sigma(t,s)\int_0^s \sigma(t,x)\,c(s,x)\,dx \tag{4.3}$$

where $c(s,y) = C_f(s,y)$.

Hence, the no-arbitrage dynamics of forward rates under the risk-neutral measure are:

$$d_t f(t,s) = \left[\frac{\partial f(t,s)}{\partial s} + \sigma(t,s)\int_0^s \sigma(t,y)\, c(s,y)\, dy\right] dt + \sigma(t,s)\, d_t Z_f(t,s) \tag{4.4}$$

Denoting the physical measure by $\mathbb{P}$ and the forward rate drift under this measure by $\mu(t,s)$, the forward rate dynamics under $\mathbb{P}$ are:

$$d_t f(t,s) = \mu(t,s)\, dt + \sigma(t,s)\, d_t Z_f^{\mathbb{P}}(t,s) \tag{4.5}$$

where $Z_f^{\mathbb{P}}(t,s)$ is a random field satisfying assumption (II) under $\mathbb{P}$. Using the results in appendix E we obtain the dynamics of bond prices under the physical measure:

$$\begin{aligned}\frac{d_t P(t,u)}{P(t,u)} &= \left[f(t,u-t) + \frac{1}{2}\int_0^{u-t}\int_0^{u-t} \sigma(t,x)\,\sigma(t,y)\, c(x,y)\, dx dy\right] dt \\ &\quad - \int_0^{u-t} \mu(t,x)\, dx dt - \int_0^{u-t} \sigma(t,x)\, d_t Z_f^{\mathbb{P}}(t,x)\, dx\end{aligned}$$

To derive the functional form of the drift $\mu(t,s)$ implied by the no-arbitrage condition, we assume the existence of a stochastic discount factor $M(t)$ with the following dynamics in each model:

$$\begin{aligned}\text{model I} &: \frac{dM(t)}{M(t)} = -r(t)\, dt - \int_0^{\infty} \lambda(t,y)\, d_t Z_f^{\mathbb{P}}(t,y)\, dy \\ \text{model II} &: \frac{dM(t)}{M(t)} = -r(t)\, dt - \int_0^{\infty} \lambda(t,y)\, d_t Z_f^{\mathbb{P}}(t,y)\, dy - \psi(t)\, dW^{\mathbb{P}}(t) \\ \text{model III} &: \frac{dM(t)}{M(t)} = -r(t)\, dt - \int_0^{\infty} \lambda(t,y)\, d_t Z_f^{\mathbb{P}}(t,y)\, dy - \int_0^{\infty} \psi(t,y)\, d_t Z_\sigma^{\mathbb{P}}(t,y)\, dy\end{aligned}$$

where $\lambda(t,s)$ denotes the forward rate risk premium, and $\psi(t)$, $\psi(t,s)$ the volatility risk premia in models II and III respectively. For no arbitrage opportunities to exist, the product $M(t)\,P(t,s)$ must be a martingale. In model III the dynamics of $P_M(t,s) = M(t)\,P(t,s)$ are:

$$\begin{aligned}\frac{d_t P_M(t,s)}{P_M(t,s)} &= -r(t)\, dt - \int_0^{u-t} \mu(t,x)\, dx dt + \int_{x=0}^{u-t}\int_{y=0}^{\infty} \sigma(t,x)\,\lambda(t,y)\, c(x,y)\, dx dy dt \\ &\quad \left[f(t,u-t) + \frac{1}{2}\int_0^{u-t}\int_0^{u-t} \sigma(t,x)\,\sigma(t,y)\, c(x,y)\, dx dy\right] dt \\ &\quad - \int_0^{\infty} \psi(t,y)\, d_t Z_\sigma^{\mathbb{P}}(t,y)\, dy - \int_0^{\infty} \lambda(t,y)\, d_t Z_f^{\mathbb{P}}(t,y)\, dy \\ &\quad - \int_0^{u-t} \sigma(t,x)\, d_t Z_f^{\mathbb{P}}(t,x)\, dx\end{aligned}$$

Hence,

$$\begin{aligned}\int_0^{u-t} \mu(t,x)\, dx &= f(t,u-t) + \frac{1}{2}\int_0^{u-t}\int_0^{u-t} \sigma(t,x)\,\sigma(t,y)\, c(x,y)\, dx dy \\ &\quad - r(t) + \int_{x=0}^{u-t}\int_{y=0}^{\infty} \sigma(t,x)\,\lambda(t,y)\, c(x,y)\, dx dy\end{aligned}$$

Let $s = u - t$. Differentiating the above equation with respect to u,

$$\mu(t,s) = \frac{\partial f(t,s)}{\partial s} + \sigma(t,s)\int_0^s \sigma(t,x)\,c(s,x)\,dx + \sigma(t,s)\int_0^\infty \lambda(t,y)\,c(s,y)\,dy$$

Hence, the forward rate drift under the physical measure in each model is as follows:

model I

$$\mu(t,s) = \frac{\partial f(t,s)}{\partial s} + g(s)\left[r(t)\int_0^s g(x)\,c(x,s)\,dx + \sqrt{r(t)}\int_0^\infty \lambda(t,u)\,c(u,s)\,du\right]$$

model II

$$\mu(t,s) = \frac{\partial f(t,s)}{\partial s} + g(s)\left[\sigma^2(t)\int_0^s g(x)\,c(x,s)\,dx + \sigma(t)\int_0^\infty \lambda(t,u)\,c(u,s)\,du\right]$$

model III

$$\mu(t,s) = \frac{\partial f(t,s)}{\partial s} + \sigma(t,s)\int_0^s \sigma(t,x)\,c(x,s)\,dx + \sigma(t,s)\int_0^\infty \lambda(t,u)\,c(u,s)\,du$$

In all models we assume that the forward rate risk premium $\lambda(t,s)$ is a function of the spot rate and the time to maturity:

$$\lambda(t,s) = r(t)\,\lambda(s) \tag{4.6}$$

where $\lambda(s)$ is a deterministic function of s.

The change of measure from the physical to the risk-neutral is given by:

$$d_t Z_f(t,s) = \int_0^\infty \lambda(t,u)\,c(u,s)\,du dt + d_t Z_f^{\mathbb{P}}(t,s) \tag{4.7}$$

To obtain the volatility dynamics under the physical measure in model II we define the process $\tilde{m}(t)$ as follows:

$$\frac{d\tilde{m}(t)}{\tilde{m}(t)} = \int_0^\infty \lambda(t,y)\,d_t Z_f(t,y)\,dy + \psi(t)\,dW(t)$$

Under the physical measure, $W^{\mathbb{P}}(t)$ defined by the following equation is a standard Brownian motion satisfying assumption (I):

$$\begin{aligned} dW^{\mathbb{P}}(t) &= dW(t) - \frac{1}{\tilde{m}(t)} d\langle W(t), \tilde{m}(t)\rangle \\ &= dW(t) - \psi(t)\,dt \end{aligned} \tag{4.8}$$

We assume that $\psi(t) = \psi$, where ψ is a constant. Hence, the dynamics of $\ln\sigma(t)$ under the physical measure in model II are:

$$d\ln\sigma(t) = \zeta\left[\vartheta - \ln\sigma(t)\right]dt + \eta\, dW^{\mathbb{P}}(t) \tag{4.9}$$

where $\vartheta = \theta + \frac{\eta\psi}{\zeta}$. Under this specification the volatility risk premium ψ is not identifiable from forward rate data alone. In this model it would make no difference if we assumed that the volatility risk premium is proportional to $\ln\sigma(t)$: $\psi(t) = \psi\ln\sigma(t)$. The dynamics of $\ln\sigma(t)$ would be of the same form:

$$d\ln\sigma(t) = \tilde{\zeta}\left[\tilde{\vartheta} - \ln\sigma(t)\right]dt + \eta dW^{\mathbb{P}}(t)$$

where $\tilde{\vartheta} = \frac{\theta\zeta}{\zeta-\eta\psi}$ and $\tilde{\zeta} = \zeta - \eta\psi$.

In model III the process $\tilde{m}(t)$ is:

$$\frac{d\tilde{m}(t)}{\tilde{m}(t)} = \int_0^\infty \lambda(t,y)\,d_tZ_f(t,y)\,dy + \int_0^\infty \psi(t,y)\,d_tZ_\sigma(t,y)\,dy$$

$Z_\sigma^{\mathbb{P}}(t,s)$ defined below is a random field satisfying assumptions (I) and (II) under the physical measure:

$$\begin{aligned} d_tZ_\sigma^{\mathbb{P}}(t,s) &= d_tZ_\sigma(t,s) - \frac{1}{\tilde{m}(t)}d\langle Z_\sigma(t,s),\tilde{m}(t)\rangle \\ &= d_tZ_\sigma(t,s) - \int_0^\infty \psi(t,y)\,C_\sigma(s,y)\,dydt \end{aligned} \tag{4.10}$$

We assume that $\psi(t,s) = \psi(s)$, where $\psi(s)$ is a deterministic function of the time to maturity s. Thus the dynamics of $\ln\sigma(t,s)$ under the physical measure in model III are:

$$d_t\ln\sigma(t,s) = \zeta(s)\left[\vartheta(s) - \ln\sigma(t,s)\right]dt + \eta(s)\,d_tZ_\sigma^{\mathbb{P}}(t,s) \tag{4.11}$$

where $\vartheta(s) = \theta(s) + \frac{\eta(s)}{\zeta(s)}\int_0^\infty \psi(y)\,C_\sigma(s,y)\,dy$. As in model II, the volatility risk premium is not identifiable from forward rate data alone.

4.3 Estimation of the models

4.3.1 Model I

In order to estimate the models we discretize the dynamics of forward rates and volatility according to an Euler scheme:[4]

$$\begin{aligned} f_{t+1,s} - f_{t,s} &= \left[\frac{\partial f_{t,s}}{\partial s} + r_tg(s)\int_0^s g(x)\,c(x,s)\,dx + (r_t)^{\frac{3}{2}}g(s)\,R(s)\right]\Delta t \\ &\quad + g(s)\sqrt{r_t}X_{t+1,s} \end{aligned} \tag{4.12}$$

where $R(s) = \int_0^\infty \lambda(u)\,c(u,s)\,du$, $X_{t,s} \sim N(0,\Delta t)$ and $cor(X_{t,s_1},X_{t,s_2}) = c(s_1,s_2)$.

Let $T+1$ denote the total number of observations and m the total number of maturities in the sample. Δt is equal to one working day, and the distance between different maturities in the sample is one month. We measure both t and s in terms of

[4] We assume that forward rates are observed without error. Furthermore, we do not model any estimation errors in the derivatives $\frac{\partial f_{t,s}}{\partial s}$.

years. Let $\mathbf{f}_i$ denote the m-dimensional column vector of forward rates at time i. The likelihood function of all $\mathbf{f}_i$, $i = 0, 1, \ldots, T$, is:

$$\begin{aligned} L(\mathbf{f}_0, \mathbf{f}_1, \ldots, \mathbf{f}_T) &= p(\mathbf{f}_0, \mathbf{f}_1, \ldots, \mathbf{f}_T) \\ &= p(\mathbf{f}_0)\, p(\mathbf{f}_1|\mathbf{f}_0)\, p(\mathbf{f}_2|\mathbf{f}_1) \cdot \ldots \cdot p(\mathbf{f}_T|\mathbf{f}_{T-1}) \end{aligned} \tag{4.13}$$

where $p(\mathbf{f}_i)$ denotes the probability density function of $\mathbf{f}_i$. Dacunha-Castelle and Florens-Zmirou (1986) show that for fixed Δt the maximum likelihood estimator of the parameters of a discretized diffusion is consistent and asymptotically normal as $T \to \infty$. Since the unconditional density of $\mathbf{f}_0$ is unknown, we maximize the *approximate* likelihood function, denoted by $\tilde{L}$:

$$\tilde{L}(\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_T) = p(\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_T|\mathbf{f}_0) \tag{4.14}$$

The omission of $p(\mathbf{f}_0)$ from the likelihood function is justified by the fact that this term is dominated by the rest of the terms in the likelihood function for large T and, as a result, does not affect results asymptotically.

Let $\mathbf{x}_t$ denote an m-dimensional column vector with jth element $x_{t,j}$ equal to:

$$x_{t,j} = \frac{f_{t+1,s_j} - f_{t,s_j} - \frac{\partial f_{t,s_j}}{\partial s}\Delta t}{g(s_j)\sqrt{r_t}} - \sqrt{r_t}\int_0^{s_j} g(x)\, c(x, s_j)\, dx \Delta t - R(s_j)\, r_t \Delta t \tag{4.15}$$

where s_j denotes the jth time to maturity, $1 \le j \le m$. Let $\boldsymbol{\Omega}$ denote the correlation matrix of the $X_{t,s}$, with ijth element equal to $c(s_i, s_j)$. We assume the following specification for the functions $g(s)$, $R(s)$, and $c(s_1, s_2)$:

$$g(s) = \alpha e^{\beta s} \;\; (\alpha > 0) \qquad R(s) = \gamma e^{\delta s} \qquad c(s_1, s_2) = e^{\kappa|s_1 - s_2|} \tag{4.16}$$

Model I is therefore described by the following vector of five parameters:

$$\boldsymbol{\theta} = (\alpha \;\; \beta \;\; \gamma \;\; \delta \;\; \kappa)' \tag{4.17}$$

The approximate log-likelihood function is equal to:

$$\begin{aligned} \ln \tilde{L} &= -\frac{1}{2\Delta t}\sum_{i=0}^{T-1} \mathbf{x}_i' \boldsymbol{\Omega}^{-1} \mathbf{x}_i - \frac{m}{2}\sum_{i=0}^{T-1} \ln r_i - T\left(\beta \sum_{j=1}^{m} s_j + m \ln a\right) \\ &\quad -\frac{T}{2}\left[m \ln(2\pi\Delta t) + \ln|\boldsymbol{\Omega}|\right] \end{aligned} \tag{4.18}$$

The maximum likelihood estimator of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \ln \tilde{L} \tag{4.19}$$

4.3.2 Model II

The discretized forward rate dynamics in model II are:

$$\begin{aligned} f_{t+1,s} - f_{t,s} &= \left[\frac{\partial f_{t,s}}{\partial s} + \sigma_t^2 g(s) \int_0^s g(x) c(x,s) \, dx + \sigma_t g(s) r_t R(s) \right] \Delta t \\ &\quad + \sigma_t g(s) X_{t+1,s} \end{aligned} \tag{4.20}$$

where $X_{t,s} \sim N(0, \Delta t)$, $R(s) = \int_0^\infty \lambda(u) c(u,s) \, du = \gamma e^{\delta s}$, and $cor(X_{t,s_1}, X_{t,s_2}) = c(s_1, s_2) = e^{\kappa|s_1 - s_2|}$, as in model I. Here $g(s)$ has the following functional form:

$$g(s) = e^{\beta s} \tag{4.21}$$

The discretized volatility dynamics are:

$$\ln \sigma_{t+1} = \ln \sigma_t + \zeta [\vartheta - \ln \sigma_t] \Delta t + \eta Y_{t+1} \tag{4.22}$$

where $Y_t \sim N(0, \Delta t)$ and $cor(X_{t_1,s}, Y_{t_2}) = 0 \ \forall \ t_1, t_2, s$. We rewrite the above equation as:

$$\ln \sigma_{t+1} = \alpha + \xi \ln \sigma_t + \phi U_{t+1} \tag{4.23}$$

where $\alpha = \zeta \vartheta \Delta t$, $\xi = 1 - \zeta \Delta t$, $\phi = \eta \sqrt{\Delta t}$, and $U_t \sim N(0,1)$. If $|\xi| < 1$ then $\ln \sigma_t$ is stationary, and the unconditional distribution of $\ln \sigma_t$ is normal with mean $\frac{\alpha}{1-\xi}$ and variance $\frac{\phi^2}{1-\xi^2}$. We assume that ξ satisfies the condition for stationarity. The distribution of $\ln \sigma_t$ conditional on $\ln \sigma_{t-1}$ is normal with mean $\alpha + \xi \ln \sigma_{t-1}$ and variance ϕ^2.

As before, we maximize the approximate likelihood function, since the unconditional density of $\mathbf{f}_0$ is unknown. The approximate likelihood function is equal to:

$$\begin{aligned} \tilde{L}(\mathbf{f}_1, \ldots, \mathbf{f}_T) &= p(\mathbf{f}_1, \ldots, \mathbf{f}_T | \mathbf{f}_0) \\ &= \int_{\mathbb{R}^T} p(\mathbf{f}_1, \ldots, \mathbf{f}_T | \mathbf{f}_0, \boldsymbol{\sigma}) \, \varphi(\boldsymbol{\sigma} | \mathbf{f}_0) \, d\boldsymbol{\sigma} \end{aligned} \tag{4.24}$$

where $\boldsymbol{\sigma}$ is a T-dimensional column vector of volatilities, $\boldsymbol{\sigma} = (\sigma_0 \ \ \sigma_1 \ \ \ldots \ \ \sigma_{T-1})'$, and $\varphi(\boldsymbol{\sigma}|\mathbf{f}_0)$ denotes the joint probability density function of $\sigma_0, \sigma_1, \ldots, \sigma_{T-1}$, given $\mathbf{f}_0$. Hence,

$$\begin{aligned} \tilde{L}(\mathbf{f}_1, \ldots, \mathbf{f}_T) &= \int_{\mathbb{R}^T} p(\mathbf{f}_1 | \mathbf{f}_0, \sigma_0) \, p(\mathbf{f}_2 | \mathbf{f}_1, \sigma_1) \, p(\mathbf{f}_3 | \mathbf{f}_2, \sigma_2) \cdot \ldots \cdot p(\mathbf{f}_T | \mathbf{f}_{T-1}, \sigma_{T-1}) \times \\ &\quad \varphi(\sigma_0 | \mathbf{f}_0) \, \varphi(\sigma_1 | \sigma_0) \cdot \ldots \cdot \varphi(\sigma_{T-1} | \sigma_{T-2}) \, d\sigma_0 d\sigma_1 \ldots d\sigma_{T-1} \end{aligned} \tag{4.25}$$

We approximate $\varphi(\sigma_0|\mathbf{f}_0)$ by the unconditional density $\varphi(\sigma_0)$, as the conditional density $\varphi(\sigma_0|\mathbf{f}_0)$ is unknown. This approximation does not affect results for large T. For simplicity we will omit $\mathbf{f}_0$ from density functions henceforth.

Let $\mathbf{F}$ denote an $m \times T$-dimensional matrix with ith column $\mathbf{f}_i$. Equation (4.25) implies that the approximate likelihood function $\tilde{L}(\mathbf{F})$ can be estimated by generating Λ volatility paths $\boldsymbol{\sigma}$ and taking the average of $p(\mathbf{F}|\boldsymbol{\sigma})$, given a forward-rate sample $(\mathbf{f}_0, \mathbf{F})$, since $\tilde{L}(\mathbf{F}) = E[p(\mathbf{F}|\boldsymbol{\sigma})]$. Hence, an estimator for the approximate likelihood is

$$\bar{L}(\mathbf{F}) = \frac{1}{\Lambda}\sum_{j=1}^{\Lambda} p\left(\mathbf{F}|\boldsymbol{\sigma}_j\right) \tag{4.26}$$

This naive Monte Carlo estimator is inefficient, however, as Λ needs to be unfeasibly large in order to obtain an accurate estimate, given a realistic sample size. There exist a number of methods that can be employed to reduce the variance of the above Monte Carlo estimator, one of which is importance sampling. In importance sampling volatility samples are not generated using $\varphi(\boldsymbol{\sigma})$, but from an alternative density, the importance sampling density $\psi(\boldsymbol{\sigma})$, which has the property that it is approximately proportional to $p(\mathbf{F}, \boldsymbol{\sigma})$:

$$\tilde{L}(\mathbf{F}) = \int \frac{p(\mathbf{F}|\boldsymbol{\sigma})\,\varphi(\boldsymbol{\sigma})}{\psi(\boldsymbol{\sigma})}\psi(\boldsymbol{\sigma})\,d\boldsymbol{\sigma} \tag{4.27}$$

The importance sampling likelihood estimator, therefore, is:

$$\hat{L}(\mathbf{F}) = \frac{1}{\Lambda}\sum_{j=1}^{\Lambda} \frac{p\left(\mathbf{F}|\boldsymbol{\sigma}_j\right)\varphi(\boldsymbol{\sigma}_j)}{\psi(\boldsymbol{\sigma}_j)} \tag{4.28}$$

Geweke (1989) shows that $\hat{L}(\mathbf{F})$ is a consistent and, under certain conditions, asymptotically normal estimator of $\tilde{L}(\mathbf{F})$. The smaller the variance of $\frac{p(\mathbf{F}|\boldsymbol{\sigma})\varphi(\boldsymbol{\sigma})}{\psi(\boldsymbol{\sigma})}$, the smaller Λ needs to be. When $\frac{p(\mathbf{F}|\boldsymbol{\sigma})\varphi(\boldsymbol{\sigma})}{\psi(\boldsymbol{\sigma})}$ is exactly equal to a constant c, then a sample of only one volatility trajectory is required. Since $p(\mathbf{F}|\boldsymbol{\sigma})\,\varphi(\boldsymbol{\sigma}) = p(\mathbf{F})\,\varphi(\boldsymbol{\sigma}|\mathbf{F})$, if $\psi(\boldsymbol{\sigma}) = \varphi(\boldsymbol{\sigma}|\mathbf{F})$ then $c = p(\mathbf{F}) = \tilde{L}(\mathbf{F})$. As $\tilde{L}(\mathbf{F})$ is the quantity which we seek to estimate, however, this result is not helpful.

Danielsson and Richard (1993), and Richard and Zhang (2000) have developed importance sampling techniques for the estimation of stochastic volatility models, where the optimal importance sampling density is obtained through an iterative process.[5] For example, Liesenfeld and Richard (2003) apply the Richard and Zhang (2000) method as follows. First a sample of volatility paths is generated from $\varphi(\boldsymbol{\sigma})$ given a set of values for the model parameters. Having selected a particular distribution, such as the normal for example, for $\psi(\sigma_t|\sigma_{t-1})$, the parameters of this density are obtained by performing a set of T regressions that seek to fit the denominator of the ratio in equation (4.28) to the numerator, using the sample of volatility paths. The regression equations are constructed

[5]Durbin and Koopman (1997) provide an alternative importance sampling method for models that can be approximated by linear Gaussian models.

by taking the logarithm of the above ratio and organizing the resulting terms in such a way that the volatility values corresponding to a single time period appear in only one regression equation. Due to this arrangement the parameters for $\psi(\sigma_{t+1}|\sigma_t)$ must be known when estimating the parameters for $\psi(\sigma_t|\sigma_{t-1})$. As a result, the parameter calculation procedure starts from the last time period. Since the volatility sample used to obtain the importance sampling density parameters was not generated using $\psi(\boldsymbol{\sigma})$, once a set of parameters for the importance sampling density is obtained, a new volatility sample is generated using these parameters, and the process is repeated to calculate new parameters, until convergence is detected in these parameter values.

Convergence is not guaranteed, however, and depending on the model it may not occur at all, as we find in our case. For this reason we propose a method for obtaining the parameters of the importance sampling density where these parameters are consistent with the sample used to obtain them. This is achieved by making the volatility sample a function of the unknown parameters. By simultaneously generating the volatility sample and calculating the parameters we eliminate the need for iterations.

We assume that $\psi(\ln\sigma_0)$ and $\psi(\ln\sigma_t|\ln\sigma_{t-1})$, $1 \leq t \leq T-1$, have the following form respectively:

$$\psi(\ln\sigma_0) = \frac{\sqrt{1-\xi^2}}{\phi\sqrt{2\pi}} \exp\left[-\frac{1-\xi^2}{2\phi^2}(\ln\sigma_0 - a_0)^2\right] \tag{4.29}$$

$$\psi(\ln\sigma_t|\ln\sigma_{t-1}) = \frac{1}{\phi\sqrt{2\pi}} \exp\left[-\frac{1}{2\phi^2}(\ln\sigma_t - a_t - \xi\ln\sigma_{t-1})^2\right] \tag{4.30}$$

where a_i, $i = 0, \ldots, T-1$, are the parameters to be obtained. Thus $\psi(\ln\sigma_0)$ differs from $\varphi(\ln\sigma_0)$ only in the mean – a_0 and $\frac{\alpha}{1-\xi}$ respectively – and $\psi(\ln\sigma_t|\ln\sigma_{t-1})$ differs from $\varphi(\ln\sigma_t|\ln\sigma_{t-1})$ only in parameter α, which is replaced by a_t. The regression equations used to obtain the parameters a_i are as follows:

$$\ln p(\mathbf{f}_1|\mathbf{f}_0, \ln\sigma_0) + \ln\varphi(\ln\sigma_0) = c_0 + \ln\psi(\ln\sigma_0) + \nu_0 \tag{4.31}$$

$$\ln p(\mathbf{f}_t|\mathbf{f}_{t-1}, \ln\sigma_{t-1}) + \ln\varphi(\ln\sigma_{t-1}|\ln\sigma_{t-2}) = c_{t-1} + \ln\psi(\ln\sigma_{t-1}|\ln\sigma_{t-2}) + \nu_{t-1} \quad \text{for } 2 \leq t \leq T \tag{4.32}$$

where the c_i, $i = 0, \ldots, T-1$, are constants, and the ν_i, $i = 0, \ldots, T-1$, are regression errors with mean zero.

We allow each σ_t to appear in two different regression equations. This is less efficient than the global optimization procedure of Liesenfeld and Richard (2003), where σ_t appears in only one regression equation, but it enables us to obtain importance sampling density parameters which are consistent with the volatility sample being used. In the

Liesenfeld and Richard (2003) procedure, the volatility sample cannot be generated simultaneously with the calculation of the importance sampling density parameters, because of the dependence of the time-t parameters on the time-$(t+1)$ parameters, which implies that the regression procedure must start from the last time period. In order to obtain a volatility sample as a function of the time-t parameters, however, the parameters for all time periods prior to t must be known. Our construction of the regression equations enables us to begin the calculation of the parameters from the first time period, so that we can simultaneously generate a volatility sample as a function of these parameters.

Let the following equation represent the regression equation for time period t:

$$y_{t,i} = c_t + z_{t,i} + \nu_{t,i} \qquad 0 \leq t \leq T-1 \tag{4.33}$$

where the subscript i denotes the ith volatility sample, $z_{t,i}$ denotes the logarithm of the importance sampling density, and $y_{t,i}$ is the sum of the logarithms of p and φ. This is not a standard non-linear regression model because both $y_{t,i}$ and $z_{t,i}$ depend on a_t through $\ln \sigma_{t,i}$. As a result, minimizing the residual sum of squares (RSS) is meaningless here. Rather, we minimize RSS divided by the total sum of squares (TSS), with RSS and TSS defined as:

$$RSS = \sum_{i=1}^{\Lambda} (y_{t,i} - c_t - z_{t,i})^2 \qquad TSS = \sum_{i=1}^{\Lambda} y_{t,i}^2 \tag{4.34}$$

where the value of the regression equation constant is:

$$\hat{c}_t = \frac{1}{\Lambda} \sum_{i=1}^{\Lambda} (y_{t,i} - z_{t,i}) \tag{4.35}$$

Replacing c_t by the above quantity, $\frac{RSS}{TSS}$ becomes a function of a_t only. The volatility sample for each time period t is a function of a_t, given the volatility sample of the previous time period if $t > 0$. $\frac{RSS}{TSS}$ is likely to have several local minima and, as a result, we minimize it numerically for each t, starting from $t = 0$. The same set of random numbers is used to generate the sample of volatility paths in each likelihood evaluation, in order to ensure the efficiency of the maximum likelihood estimation procedure.[6] This set of random numbers consists of T subsets, one for each time period t. Each subset is a sample of Λ numbers from a standard normal distribution and is used to obtain the time-t volatility sample employed in the evaluation of $\frac{RSS}{TSS}$ during the numerical search for the optimal value of the parameter a_t, as well as the final volatility sample for time t used in the estimation of the likelihood function given the optimal value for a_t.

[6] See Gourieroux and Monfort (1997).

Let $\mathbf{x}_t$ denote an m-dimensional column vector with jth element $x_{t,j}$ equal to:

$$x_{t,j} = \frac{f_{t+1,s_j} - f_{t,s_j} - \frac{\partial f_{t,s_j}}{\partial s}\Delta t}{e^{\beta s_j}\sigma_t} - \sigma_t \int_0^{s_j} e^{\beta x} c(x, s_j)\, dx \Delta t - r_t R(s_j)\Delta t \tag{4.36}$$

where s_j denotes the jth time to maturity, $1 \leq j \leq m$. As before, $\boldsymbol{\Omega}$ denotes the correlation matrix of the $X_{t,s}$.

The variables $y_{t,i}$ and $z_{t,i}$ defined earlier are given by the following equations:

$$y_{t,i} = \frac{1}{2\Delta t}\mathbf{x}_{t,i}'\boldsymbol{\Omega}^{-1}\mathbf{x}_{t,i} + m\ln\sigma_{t,i} + \frac{1}{2\phi^2}(\ln\sigma_{t,i} - \xi\ln\sigma_{t-1,i} - \alpha)^2 \tag{4.37}$$

$$z_{t,i} = \frac{1}{2\phi^2}(\ln\sigma_{t,i} - \xi\ln\sigma_{t-1,i} - a_t)^2 \tag{4.38}$$

for $1 \leq t \leq T-1$ and

$$y_{0,i} = \frac{1}{2\Delta t}\mathbf{x}_{0,i}'\boldsymbol{\Omega}^{-1}\mathbf{x}_{0,i} + m\ln\sigma_{0,i} + \frac{1-\xi^2}{2\phi^2}\left(\ln\sigma_{0,i} - \frac{\alpha}{1-\xi}\right)^2 \tag{4.39}$$

$$z_{0,i} = \frac{1-\xi^2}{2\phi^2}(\ln\sigma_{0,i} - a_0)^2 \tag{4.40}$$

where $\mathbf{x}_{t,i}$ denotes the value of the vector $\mathbf{x}_t$ given the ith volatility sample.

Model II is described by a vector of seven parameters:

$$\boldsymbol{\theta} = (\beta \quad \gamma \quad \delta \quad \kappa \quad \alpha \quad \xi \quad \phi)' \tag{4.41}$$

Let

$$\begin{aligned} q_i = & -\frac{1}{2\Delta t}\sum_{t=0}^{T-1}\mathbf{x}_{t,i}'\boldsymbol{\Omega}^{-1}\mathbf{x}_{t,i} - m\sum_{t=0}^{T-1}\ln\sigma_{t,i} - T\beta\sum_{j=1}^{m}s_j \\ & -\frac{T}{2}\left[\ln|\boldsymbol{\Omega}| + m\ln(2\pi\Delta t)\right] - \frac{1-\xi^2}{2\phi^2}\left(\ln\sigma_{0,i} - \frac{\alpha}{1-\xi}\right)^2 \\ & -\frac{1}{2\phi^2}\sum_{t=1}^{T-1}(\ln\sigma_{t,i} - \xi\ln\sigma_{t-1,i} - \alpha)^2 + \frac{1-\xi^2}{2\phi^2}(\ln\sigma_{0,i} - a_0)^2 \\ & +\frac{1}{2\phi^2}\sum_{t=1}^{T-1}(\ln\sigma_{t,i} - \xi\ln\sigma_{t-1,i} - a_t)^2 \end{aligned} \tag{4.42}$$

The approximate log-likelihood function is:

$$\ln\tilde{L} = \ln\left(\frac{1}{\Lambda}\sum_{i=1}^{\Lambda}e^{q_i}\right) \tag{4.43}$$

4.3.3 Model III

The discretized forward rate dynamics are:

$$\begin{aligned} f_{t+1,s} - f_{t,s} = & \left[\frac{\partial f_{t,s}}{\partial s} + \sigma_{t,s}\int_0^s \sigma_{t,x}c(x,s)\,dx + \sigma_{t,s}r_t R(s)\right]\Delta t \\ & +\sigma_{t,s}X_{t+1,s} \end{aligned} \tag{4.44}$$

As before, $X_{t,s} \sim N(0,\Delta t)$, $R(s) = \int_0^\infty \lambda(u)\, c(u,s)\, du = \gamma e^{\delta s}$, and $cor(X_{t,s_1}, X_{t,s_2}) = c(s_1,s_2) = e^{\kappa|s_1-s_2|}$.

The discretized volatility dynamics are:

$$\ln\sigma_{t+1,s} = \ln\sigma_{t,s} + \zeta_s\left[\vartheta_s - \ln\sigma_{t,s}\right]\Delta t + \eta_s Y_{t+1,s} \tag{4.45}$$

where $Y_{t,s} \sim N(0,\Delta t)$ and $cor(Y_{t,s}, Y_{t,u}) = \rho(s,u)$. We rewrite the above dynamics as:

$$\ln\sigma_{t+1,s} = \alpha_s + \xi_s \ln\sigma_{t,s} + \phi_s U_{t+1,s} \tag{4.46}$$

where $\alpha_s = \zeta_s\theta_s\Delta t$, $\xi_s = 1 - \zeta_s\Delta t$, $\phi_s = \eta_s\sqrt{\Delta t}$, $U_{t,s} \sim N(0,1)$, $cor(U_{t,s_1}, U_{t,s_2}) = \rho(s_1,s_2)$. The distribution of $\ln\sigma_{t,s}$ conditional on $\ln\sigma_{t-1,s}$ is $N\left(\alpha_s + \xi_s\ln\sigma_{t-1,s}, \phi_s^2\right)$. If $|\xi_s| < 1$ then $\ln\sigma_{t,s}$ is stationary with unconditional distribution $N\left(\frac{\alpha_s}{1-\xi_s}, \frac{\phi_s^2}{1-\xi_s^2}\right)$. We assume that ξ_s satisfies the condition for stationarity for all s. In this case the unconditional correlation between $\ln\sigma_{t,s_1}$ and $\ln\sigma_{t,s_2}$ is equal to:

$$cor(\ln\sigma_{t,s_1}, \ln\sigma_{t,s_2}) = \frac{\sqrt{\left(1-\xi_{s_1}^2\right)\left(1-\xi_{s_2}^2\right)}}{1-\xi_{s_1}\xi_{s_2}}\rho(s_1,s_2) \tag{4.47}$$

We assume the following functional forms for $\rho(x,y)$, α_s, ξ_s, and ϕ_s:

$$\rho(x,y) = e^{\rho|x-y|} \qquad \alpha_s = \alpha_0 e^{\alpha s} \qquad \xi_s = \xi_0 e^{\xi s} \qquad \phi_s = \phi_0 e^{\phi s} \tag{4.48}$$

where ρ, α, α_0, ξ, ξ_0, ϕ, and ϕ_0 are constants.

Let $\mathbf{F}$ be defined as in subsection 4.3.2, and $\boldsymbol{\sigma}_t$ be an m-dimensional column vector of volatilities with ith element equal to $\ln\sigma_{t,s_i}$. Furthermore, let $\mathbf{x}_t$ denote an m-dimensional column vector with jth element $x_{t,j}$ equal to

$$x_{t,j} = \frac{f_{t+1,s_j} - f_{t,s_j} - \frac{\partial f_{t,s_j}}{\partial s}\Delta t}{\sigma_{t,s_j}} - \int_0^{s_j} \sigma_{t,x} c(x,s_j)\, dx\Delta t - r_t R(s_j)\Delta t \tag{4.49}$$

where s_j denotes the jth time to maturity, $1 \le j \le m$. We approximate the integral in the above equation by the following sum:

$$\int_0^{s_j} \sigma_{t,x} c(x,s_j)\, dx \approx \sum_{i=1}^{j} \sigma_{t,s_i} c(s_i,s_j)\Delta s - \frac{1}{2}\left[\sigma_{t,s_j} + \sigma_{t,s_1} c(s_1,s_j)\right]\Delta s \tag{4.50}$$

where $\Delta s = \frac{1}{12}$ corresponding to increments of one month.

Let $\mathbf{u}_t$ denote an m-dimensional column vector with jth element $u_{t,j}$ equal to

$$u_{0,j} = \frac{\sqrt{1-\xi_0^2 e^{2\xi s_j}}}{\phi_0 e^{\phi s_j}}\left(\ln\sigma_{t,s_j} - \frac{\alpha_0 e^{\alpha s_j}}{1-\xi_0 e^{\xi s_j}}\right) \quad \text{for} \quad t = 0 \tag{4.51}$$

$$u_{t,j} = \frac{1}{\phi_0 e^{\phi s_j}}\left(\ln\sigma_{t,s_j} - \alpha_0 e^{\alpha s_j} - \xi_0 e^{\xi s_j}\ln\sigma_{t-1,s_j}\right) \quad \text{for} \quad 1 \le t \le T-1 \tag{4.52}$$

In order to estimate the likelihood function we adapt the importance sampling method used for model II. The unconditional and conditional volatility densities, $\varphi(\boldsymbol{\sigma}_0)$ and $\varphi(\boldsymbol{\sigma}_t|\boldsymbol{\sigma}_{t-1})$, are:

$$\varphi(\boldsymbol{\sigma}_0) = \frac{1}{(2\pi)^{\frac{m}{2}}|\boldsymbol{\Sigma}_{unc}|^{\frac{1}{2}}\phi_0^m\tilde{\phi}_0}\exp\left[-\frac{1}{2}\mathbf{u}_0'\boldsymbol{\Sigma}_{unc}^{-1}\mathbf{u}_0\right] \tag{4.53}$$

$$\varphi(\boldsymbol{\sigma}_t|\boldsymbol{\sigma}_{t-1}) = \frac{1}{(2\pi)^{\frac{m}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}\phi_0^m\tilde{\phi}}\exp\left[-\frac{1}{2}\mathbf{u}_t'\boldsymbol{\Sigma}^{-1}\mathbf{u}_t\right] \tag{4.54}$$

respectively, where

$$\tilde{\phi}_0 = \exp\left(\phi\sum_{j=1}^{m}s_j\right)\prod_{j=1}^{m}\frac{1}{\sqrt{1-\xi_0^2e^{2\xi s_j}}} \quad \text{and} \quad \tilde{\phi} = \exp\left(\phi\sum_{j=1}^{m}s_j\right) \tag{4.55}$$

and $\boldsymbol{\Sigma}$, $\boldsymbol{\Sigma}_{unc}$ are the conditional and unconditional correlation matrices of $\ln\sigma_{t,s}$ respectively.

We choose importance sampling densities $\psi(\boldsymbol{\sigma}_0)$ and $\psi(\boldsymbol{\sigma}_t|\boldsymbol{\sigma}_{t-1})$ which differ from $\varphi(\boldsymbol{\sigma}_0)$ and $\varphi(\boldsymbol{\sigma}_t|\boldsymbol{\sigma}_{t-1})$ respectively only in parameter α_0, which is replaced by a_0 and a_t respectively. Let $\mathbf{w}_t$ denote a vector which is equal to $\mathbf{u}_t$ for $\alpha_0 = a_t$. The unconditional and conditional importance sampling densities are given by the following two formulae respectively:

$$\psi(\boldsymbol{\sigma}_0) = \frac{1}{(2\pi)^{\frac{m}{2}}|\boldsymbol{\Sigma}_{unc}|^{\frac{1}{2}}\phi_0^m\tilde{\phi}_0}\exp\left[-\frac{1}{2}\mathbf{w}_0'\boldsymbol{\Sigma}_{unc}^{-1}\mathbf{w}_0\right] \tag{4.56}$$

$$\psi(\boldsymbol{\sigma}_t|\boldsymbol{\sigma}_{t-1}) = \frac{1}{(2\pi)^{\frac{m}{2}}|\boldsymbol{\Sigma}|^{\frac{1}{2}}\phi_0^m\tilde{\phi}}\exp\left[-\frac{1}{2}\mathbf{w}_t'\boldsymbol{\Sigma}^{-1}\mathbf{w}_t\right] \tag{4.57}$$

We obtain the parameters a_t, $0 \le t \le T-1$, by minimising the ratio $\frac{RSS}{TSS}$ corresponding to the following regression equation for time period t:

$$y_{t,i} = c_t + z_{t,i} + \nu_{t,i} \tag{4.58}$$

where i denotes the ith volatility sample. The variables $y_{t,i}$ and $z_{t,i}$ are given by:

$$y_{t,i} = \frac{1}{2\Delta t}\mathbf{x}_{t,i}'\boldsymbol{\Omega}^{-1}\mathbf{x}_{t,i} + \sum_{j=1}^{m}\ln\sigma_{t,s_j,i} + \frac{1}{2}\mathbf{u}_{t,i}'\boldsymbol{\Sigma}^{-1}\mathbf{u}_{t,i} \tag{4.59}$$

$$z_{t,i} = \frac{1}{2}\mathbf{w}_{t,i}'\boldsymbol{\Sigma}^{-1}\mathbf{w}_{t,i} \tag{4.60}$$

for $1 \le t \le T-1$ and

$$y_{0,i} = \frac{1}{2\Delta t}\mathbf{x}_{0,i}'\boldsymbol{\Omega}^{-1}\mathbf{x}_{0,i} + \sum_{j=1}^{m}\ln\sigma_{0,s_j,i} + \frac{1}{2}\mathbf{u}_{0,i}'\boldsymbol{\Sigma}_{unc}^{-1}\mathbf{u}_{0,i} \tag{4.61}$$

$$z_{0,i} = \frac{1}{2}\mathbf{w}_{0,i}'\boldsymbol{\Sigma}_{unc}^{-1}\mathbf{w}_{0,i} \tag{4.62}$$

for $t = 0$. As before, $\boldsymbol{\Omega}$ denotes the correlation matrix of the $X_{t,s}$, and the optimal value of the constant c_t is given by equation (4.35). We use a constant set of random numbers to generate volatility samples in order to obtain the parameters a_t and estimate the likelihood function, as described in the subsection for model II.

Model III is described by a vector of ten parameters:

$$\boldsymbol{\theta} = (\gamma \quad \delta \quad \kappa \quad \rho \quad \alpha_0 \quad \alpha \quad \xi_0 \quad \xi \quad \phi_0 \quad \phi)' \tag{4.63}$$

Let

$$\begin{aligned} q_i \quad = \quad & -\frac{1}{2\Delta t}\sum_{t=0}^{T-1}\mathbf{x}_{t,i}'\boldsymbol{\Omega}^{-1}\mathbf{x}_{t,i} - \sum_{t=0}^{T-1}\sum_{j=1}^{m}\ln\sigma_{t,s_j,i} - \frac{T}{2}\left[\ln|\boldsymbol{\Omega}| + m\ln(2\pi\Delta t)\right] \\ & -\frac{1}{2}\mathbf{u}_{0,i}'\boldsymbol{\Sigma}_{unc}^{-1}\mathbf{u}_{0,i} - \frac{1}{2}\sum_{t=1}^{T-1}\mathbf{u}_{t,i}'\boldsymbol{\Sigma}^{-1}\mathbf{u}_{t,i} + \frac{1}{2}\mathbf{w}_{0,i}'\boldsymbol{\Sigma}_{unc}^{-1}\mathbf{w}_{0,i} \\ & +\frac{1}{2}\sum_{t=1}^{T-1}\mathbf{w}_{t,i}'\boldsymbol{\Sigma}^{-1}\mathbf{w}_{t,i} \end{aligned} \tag{4.64}$$

The approximate log-likelihood function is:

$$\ln\tilde{L} = \ln\left(\frac{1}{\Lambda}\sum_{i=1}^{\Lambda}e^{q_i}\right) \tag{4.65}$$

In all three models the log-likelihood function is maximized numerically with respect to $\boldsymbol{\theta}$ using a global optimization algorithm, since $\tilde{L}(\boldsymbol{\theta})$ is likely to have many local maxima.

4.4 The data

We use two datasets which consist of daily UK and US yields obtained from Datastream. Each data point is the annual yield on a zero-coupon bond, with time to maturity ranging from zero to 120 months, at increments of one month. We use as many maturities as were available in order to approximate the continuous nature of random fields as closely as possible. The data span the time period 1/4/1997–30/1/2004. The starting date of the sample is dictated by the lack of data for long maturities at earlier dates. We measure t and s in terms of years. Thus $\Delta s = \frac{1}{12}$ and Δt is equal to the timespan of the sample ($6\frac{5}{6}$ years) divided by the number of observations, or working days, in the sample. For the UK the total number of observations is 1727, while for the US it is 1732.

Denoting yields by $y(t,s)$, the following relationship holds between yields and forward rates:

$$sy(t,s) = \int_0^s f(t,x)\,dx \tag{4.66}$$

Hence, $f(t,s)$ is equal to the partial derivative of the above quantity with respect to s. We calculate this derivative numerically by fitting a cubic spline to the yield curve data multiplied by the corresponding time to maturity. We also need the partial derivatives of forward rates with respect to s in order to estimate the models. We obtain these derivatives numerically as well, by fitting a cubic spline to each day's forward rate data.

4.5 Empirical results

The maximization of the likelihood function is performed using the simulated annealing global optimization algorithm by Press *et al.* (2002). All numerical procedures are implemented using the C++ programming language on a Solaris 9 UNIX workstation. In the case of model I, which is not as computationally demanding as the other models, we also verified the results by a grid search of 3000 points.

Tables 4.2 – 4.5 present the parameter estimates (to three significant figures) and the maximum of the log-likelihood function for each model. Appendix F presents the graphs of the exponential functions $g(s)$, $R(s)$, $c(s_1 - s_2)$, $\alpha(s)$, $\xi(s)$, $\phi(s)$, and $\rho(s_1 - s_2)$ given the estimated values of the parameters of each model.

	α	β	γ	δ	κ
UK	0.232	−0.181	-3.36	0.113	−0.381
US	0.151	−0.0513	-6.15	−0.0411	-1.16

Table 4.2: Model I parameter estimates

	β	γ	δ	κ	α	ξ	ϕ
UK	−0.136	-9.28	0.218	−0.259	−0.962	0.318	0.402
US	0.0115	-8.30	0.379	−1.14	−1.36	0.380	0.337

Table 4.3: Model II parameter estimates

Tables 4.6 – 4.8 present the standard errors of the parameter estimates, calculated numerically.

	γ	δ	κ	α_0	α
UK	–0.00291	–0.761	–0.000300	–0.202	–0.236
US	–8.88	–0.559	–0.000655	–0.341	–0.0337
	ρ	ξ_0	ξ	ϕ_0	ϕ
UK	–1.23	0.585	–0.278	0.369	–0.236
US	–1.029	0.0694	–0.0762	0.571	–0.0337

Table 4.4: Model III parameter estimates

	Model I	Model II	Model III
UK	1678760	1376580	1301290
US	1614620	1285940	1250170

Table 4.5: Log-likelihood

	α	β	γ	δ	κ
UK	0.000359	0.000243	1.87	0.065	0.00121
US	0.000233	0.000257	3.80	0.119	0.00378

Table 4.6: Model I standard errors of parameter estimates

	β	γ	δ	κ
UK	0.0000174	0.00825	0.000192	0.0000257
US	0.00000242	0.00172	0.0000551	0.000145
	α	ξ	ϕ	
UK	0.000190	0.0000290	0.0000486	
US	0.000200	0.0000264	0.0000804	

Table 4.7: Model II standard errors of parameter estimates

	γ	δ	κ	α_0	α
UK	0.00000302	0.00281	0.000000750	0.000572	0.00000819
US	0.0169	0.00216	0.000000785	0.000751	0.0000273
	ρ	ξ_0	ξ	ϕ_0	ϕ
UK	0.00470	0.00131	0.00107	0.00129	0.00000777
US	0.00129	0.000226	0.000223	0.000626	0.0000230

Table 4.8: Model III standard errors of parameter estimates

	φ_0	φ_1	σ
true values	0	0.97	0.2
mean	0.00330	0.873	0.205
st. deviation	0.0522	0.169	0.0767
skewness	3.42	-3.49	3.23
kurtosis	19.0	13.7	13.5

Table 4.9: Monte Carlo study

Table 4.9 presents a Monte Carlo study of our importance sampling technique for estimating stochastic volatility models using Monte Carlo maximum likelihood. We estimate the following simple stochastic volatility model:

$$y_t = \sqrt{x_t} U_t \tag{4.67}$$

$$\ln x_t = \varphi_0 + \varphi_1 \ln x_{t-1} + \sigma Z_t \tag{4.68}$$

where U_t and Z_t are independent standard normal random variables. We generate a sample of size 1000 given the model parameter values shown in the second row of the above table. Using the sample for y_t we estimate the model 1000 times. We summarize our results in rows three to six of the above table. The importance sampling density differs from the actual density of $\ln x_t$ only in the mean and we obtain the optimal value of the importance sampling density mean in the way described in section 4.3. Appendix G presents the graphs of the distributions of the obtained estimates for parameters φ_0, φ_1, and σ.

Let q denote the average of the ratio $\frac{RSS}{TSS}$ over all time periods obtained in one likelihood evaluation for model II or III:

$$q = \frac{1}{T} \sum_{t=0}^{T-1} \frac{RSS_t}{TSS_t} \tag{4.69}$$

For model II the average value of q over all likelihood evaluations performed is 0.63% for the UK and 2.18% for the US. The standard deviations are 1.35% and 5.84%, and the total number of likelihood evaluations are 2110 and 4000 respectively. For model III the average value of q is 0.9% for the UK and 1.6% for the US. The standard deviations are 6.8% and 4.9%, and the total number of likelihood evaluations are 2980 and 2200 respectively.

However, q is not a reliable measure of the fit of the importance sampling density to the joint density of forward rates and volatility, since we allow $\ln \sigma_t$ to appear in two different regression equations, as explained in subsection 4.3.2. Let $y_j(\boldsymbol{\theta})$ denote the value of the joint density of forward rates and volatility obtained for parameter vector $\boldsymbol{\theta}$ and the jth volatility sample given $\boldsymbol{\theta}$, and let $z_j(\boldsymbol{\theta})$ denote the corresponding value of the importance sampling density. Denoting the total number of volatility samples used in each likelihood evaluation by Λ, our estimate of the likelihood for parameter vector $\boldsymbol{\theta}$ is equal to the following mean:

$$\tilde{L}(\boldsymbol{\theta}) = \frac{1}{\Lambda}\sum_{j=1}^{\Lambda}\frac{y_j(\boldsymbol{\theta})}{z_j(\boldsymbol{\theta})} \tag{4.70}$$

Let

$$c(\boldsymbol{\theta}) = \frac{1}{\Lambda}\sum_{j=1}^{\Lambda}\left(\ln y_j(\boldsymbol{\theta}) - \ln z_j(\boldsymbol{\theta})\right) \tag{4.71}$$

A measure of the fit of the importance sampling density to the joint density of forward rates and volatility is given by the following quantity:

$$\tilde{r}(\boldsymbol{\theta}) = \sum_{j=1}^{\Lambda}\left[\ln y_j(\boldsymbol{\theta}) - \ln z_j(\boldsymbol{\theta}) - c(\boldsymbol{\theta})\right]^2 / \sum_{j=1}^{\Lambda}\left[\ln y_j(\boldsymbol{\theta})\right]^2 \tag{4.72}$$

The above quantity is equal to the residual sum of squares divided by the total sum of squares for the following model:

$$\ln y_j(\boldsymbol{\theta}) = \ln z_j(\boldsymbol{\theta}) + c(\boldsymbol{\theta}) + \text{error} \tag{4.73}$$

For model II the average value of $\tilde{r}$ over all likelihood evaluations is 0.19% for the UK and 1.14% for the US. The standard deviations are 3.44% and 9.18% respectively. However, apart from a few large outliers, most values of $\tilde{r}$ (98.6% and 96.4% for the UK and the US respectively) fall below 0.1% for $\Lambda = 20$. Hence, our importance sampling method provides a highly accurate likelihood estimate. For model III the average value of $\tilde{r}$ is 0.61% for the UK and 3% for the US. The standard deviations are 6.4% and 15% respectively, while 98.8% and 96% of the values of $\tilde{r}$ fall below 0.1% for the UK and the US respectively for $\Lambda = 20$.

4.6 Conclusion

Not much empirical research exists concerning multifactor term structure models. Previous econometric studies of stochastic volatility models have concentrated on models with a low number of latent variables, due to the estimation difficulties posed by high-dimensional models. In this chapter we estimate random field models of the term structure of interest rates using 121 forward rate maturities and up to 121 latent variables representing stochastic forward rate volatility. We use three different specifications for the forward rate volatility function: (i) a function of the spot rate and the time to maturity, (ii) a function of the time to maturity and a stochastic process driven by a Brownian motion, (iii) a random field. In cases (ii) and (iii) we assume that the random shocks driving volatility are uncorrelated with those of forward rates. The deterministic volatility model is estimated by maximum likelihood and the two stochastic volatility models by Monte Carlo maximum likelihood. We develop an importance sampling technique that substantially reduces the variance of the Monte Carlo likelihood estimator so that accurate likelihood estimates can be obtained in spite of the size of the model.

Possible avenues for future empirical research are the examination of alternative specifications for the random field models, such as the incorporation of jumps in forward rates and/or volatility, and allowing for correlation between the random field driving forward rates and innovations in volatility; the comparison of random field models with other types of term structure models; the use of option price data to test the models and estimate the volatility risk premium; and the estimation and prediction of volatility. Moreover, there is a need for further research on econometric methods for estimating high-dimensional stochastic volatility models.[7]

[7] An implementation of MCMC was found to require more computer time than Monte Carlo maximum likelihood estimation, while it exhibited very poor convergence (slow mixing) in the estimated parameter and volatility values. Even in the case of model I, which is the simplest, MCMC takes much longer than maximum likelihood estimation. Due to the complexity of the model, Metropolis-Hastings algorithms need to be used for sampling at least three of the five parameters of model I. This implies that the likelihood function needs to be evaluated more times in MCMC than in maximum likelihood, in order to obtain a large enough sample of parameter values. In the MCMC estimation of models II and III individual likelihood evaluations are as fast as in model I, but the number of simulations has to be much larger due to the inclusion of volatility in the estimated parameters.

Chapter 5

Conclusion

In this thesis we have considered random field models for the implied volatility of stock options and the term structure of interest rates. Random field models offer greater flexibility in fitting observed asset prices than previously studied finite-factor models. This advantage can eliminate the need for periodic recalibration of the model as new market data become available. It also allows greater precision in modelling the correlation between assets with similar characteristics, which can be very valuable in forming hedging portfolios for exotic derivatives using standard options or bonds.

In general, finite-factor models with n factors predict that any derivative asset can be perfectly hedged using any n bonds in term structure models, or n plain-vanilla European options in stochastic volatility stock price models, whereas in practice derivatives tend to be hedged with assets of similar maturity. While in principle, the entire continuum of bonds or options, depending on the model, is required to hedge a derivative instrument in random field models, it may be possible to attain a near-perfect hedge using only a few assets, if those assets are strongly correlated with the derivative. In the case of interest rate contingent claims it is usually easy to select those bonds which are likely to be highly correlated with the derivative instrument. Selecting options to hedge other derivatives may be less easy, however, as options have an additional exercise price dimension. In this case, the appropriate options may be chosen on the basis of liquidity issues.

Although random field models are infinite-factor models, it is still possible to obtain closed-form solutions for the prices of derivative assets, as we do in chapter 2 for variance swaptions and in chapter 3 for standard interest rate contingent claims. Furthermore, random field models can be estimated using a small number of parameters. In chapter 4 we estimate three random field models of the term structure of interest rates, with deterministic and stochastic forward rate volatility. We use 121 forward rate maturities in order to approximate as closely as possible the continuous nature of random field models using available data. Previous empirical studies of random field models in finance use a very small number of factors in the estimation process, which makes the estimated models indistinguishable from finite-factor ones. This reduction of random field models to finite-factor versions was considered necessary in order to make the estimation problem computationally feasible. By developing an appropriate importance sampling method, we are able to reduce the computational demands of the problem and estimate the stochastic volatility models using Monte Carlo maximum likelihood. There exists substantial scope for further theoretical and empirical research in the area of random field asset pricing models.

Bibliography

Ahn D.H., Dittmar R.F., Gallant A.R., Gao B., 2003, Purebred or Hybrid?: Reproducing the Volatility in Term Structure Dynamics, *Journal of Econometrics*, 116, 147–180.

Ait-Sahalia Y., 2002, Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-Form Approximation Approach, *Econometrica*, 70, 223–262.

Ait-Sahalia Y., Kimmel R., 2004, Maximum Likelihood Estimation of Stochastic Volatility Models, Working Paper, Princeton University.

Amin K., Morton A.J., 1994, Implied Volatility Functions in Arbitrage-Free Term Structure Models, *Journal of Financial Economics*, 35, 141–180.

Amin K., Ng V.K., 1993, Option Valuation with Systematic Stochastic Volatility, *Journal of Finance*, 48, 881–910.

Andersen L., Andreasen J., 2002, Volatile Volatilities, *Risk*, 15, December, 163–168.

Andersen L., Brotherton-Ratcliffe R., 2001, Extended LIBOR Market Models with Stochastic Volatility, Working Paper, Gen Re Securities.

Andersen T.G., 1994, Stochastic Autoregressive Volatility: A Framework for Volatility Modeling, *Mathematical Finance*, 4, 75–102.

Andersen T.G., Lund J., 1997, Estimating Continuous-Time Stochastic Volatility Models of the Short-Term Interest Rate, *Journal of Econometrics*, 77, 343–377.

Andersen T.G., Sorensen B., 1996, GMM Estimation of a Stochastic Volatility Model: a Monte Carlo Study, *Journal of Business and Economic Statistics*, 14, 328–352.

Andreasen J., Collin-Dufresne P., Shi W., 1998, Applying the HJM Approach when Volatility is Stochastic, Working Paper.

Bailey W., Stulz R.M., 1989, The Pricing of Stock Index Options in a General Equilibrium Model, *Journal of Financial and Quantitative Analysis*, 24, 1–12.

Bajeux-Besnainou I., Rochet J.C., 1996, Dynamic Spanning: Are Options an Appropriate Instrument?, *Mathematical Finance*, 6, 1–16.

Bakshi G.S., Cao C., Chen Z., 1997, Empirical Performance of Alternative Option

Pricing Models, *Journal of Finance*, 52, 2003–2049.

Bakshi G.S., Cao C., Chen Z., 2000, Do Call Prices and the Underlying Stock Always Move in the Same Direction?, *Review of Financial Studies*, 13, 549–584.

Bakshi G.S., Chen Z., 1997, An Alternative Valuation Model for Contingent Claims, *Journal of Financial Economics*, 44, 123–165.

Bakshi G.S., Kapadia N., 2003, Delta-Hedged Gains and the Negative Market Volatility Risk Premium, *Review of Financial Studies*, 16, 527–566.

Ball C.A., Torous W.N., 1999, The Stochastic Volatility of Short-Term Interest Rates: Some International Evidence, *Journal of Finance*, 54, 2339–2359.

Bates D.S., 1997, The Skewness Premium: Option Pricing Under Asymmetric Processes, *Advances in Futures and Options Research*, 9, 51–82.

Beckers S., 1980, The Constant Elasticity of Variance Model and its Implications for Option Pricing, *Journal of Finance*, 35, 661–673.

Bensoussan A., Crouhy M., Galai D., 1994, Stochastic Equity Volatility and the Capital Structure of the Firm, *Philosophical Transactions: Physical Sciences and Engineering*, 347, 531–540.

Benzoni L., 2002, Pricing Options under Stochastic Volatility: An Empirical Investigation, Working Paper.

Black F., 1976, The Pricing of Commodity Contracts, *Journal of Financial Economics*, 3, 167–179.

Black F., Scholes M., 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*, 81, 637–654.

Blattberg R.C., Gonedes N.J., 1974, A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices, *Journal of Business*, 47, 244–280.

Bollerslev T., Chou R.Y., Kroner K.F., 1992, ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence, *Journal of Econometrics*, 52, 5–59.

Brace A., Gatarek D., Musiela M., 1997, The Market Model of Interest Rate Dynamics, *Mathematical Finance*, 7, 127–155.

Brace A., Goldys B., Klebaner F., Womersley R., 2001, Market Model of Stochastic Implied Volatility with Application to the BGM Model, Working Paper S01–1, Department of Statistics, University of New South Wales.

Brace A., Goldys B., Van Der Hoek J., Womersley R., 2002, Markovian Models in the Stochastic Implied Volatility Framework, Working Paper.

Brandt M.W., Santa-Clara P., 2002, Simulated Likelihood Estimation of Diffusions with an Application to Exchange Rate Dynamics in Incomplete Markets, *Journal* *of Financial Economics*, 63, 161–210.

Brennan M.J., Schwartz E.S., 1979, A Continuous Time Approach to the Pricing of Bonds, *Journal of Banking and Finance*, 3, 133–155.

Britten-Jones M., Neuberger A., 2000, Option Prices, Implied Price Processes and Stochastic Volatility, *Journal of Finance*, 55, 839–866.

Buraschi A., Jackwerth J., 2001, The Price of a Smile: Hedging and Spanning in Option Markets, *Review of Financial Studies*, 14, 495–527.

Carrasco M., Chernov M., Florens J.-P., Ghysels E., 2004, Efficient Estimation of Jump Diffusions and General Dynamic Models with a Continuum of Moment Conditions, Working Paper.

Carrasco M., Florens J.-P., 2002, Efficient GMM Estimation Using the Empirical Characteristic Function, Working Paper.

Chacko G., Viceira L.M., 2003, Spectral GMM Estimation of Continuous-Time Processes, *Journal of Econometrics*, 116, 259–292.

Chen L., 1996, Stochastic Mean and Stochastic Volatility – A Three-Factor Model of the Term Structure of Interest Rates and Its Applications in Derivatives Pricing and Risk Management, *Financial Markets, Institutions and Instruments*, 5, 1–87.

Chiarella C., Kwon O.K., 1999, A Class of Heath-Jarrow-Morton Term Structure Models with Stochastic Volatility, Working Paper, School of Finance and Economics, University of Technology, Sydney.

Chiarella C., Pasquali S., Runggaldier W.J., 2001, On Filtering in Markovian Term Structure Models (an Approximation Approach), Working Paper.

Chib S., Nardari F., Shephard N., 2002, Analysis of High Dimensional Multivariate Stochastic Volatility Models, Working Paper.

Chu M., 1996, The Random Yield Curve and Interest Rate Options, Working Paper, Centre for Quantitative Finance, Imperial College, London.

Collin-Dufresne P., Goldstein R., 2002, Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility, *Journal of Finance*, 57, 1685–1730.

Collin-Dufresne P., Goldstein R., 2003, Generalizing the Affine Framework to HJM and Random Field Models, Working Paper.

Collin-Dufresne P., Goldstein R., Jones C., 2003, Identification and Estimation of 'Maximal' Affine Term Structure Models: An Application to Stochastic Volatility, Working Paper.

Comte F., Renault E., 1998, Long Memory in Continuous-Time Stochastic Volatility

Models, *Mathematical Finance*, 8, 291–323.

Cox J.C., Ingersoll J.E., Ross S.A., 1985, A Theory of the Term Structure of Interest Rates, *Econometrica*, 53, 385–408.

Cox J.C., Ross S.A., 1976, The Valuation of Options for Alternative Stochastic Processes, *Journal of Financial Economics*, 3, 145–166.

Da Prato G., Zabczyk J., 1992, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge.

Dacunha-Castelle D., Florens-Zmirou D., 1986, Estimation of the Coefficients of a Diffusion from Discrete Observations, *Stochastics*, 19, 263–284.

Dai Q., Singleton K., 2003, Term Structure Dynamics in Theory and Reality, *Review of Financial Studies*, 16, 631–678.

Danielsson J., Richard J.F., 1993, Accelerated Gaussian Importance Sampler with Application to Dynamic Latent Variable Models, *Journal of Applied Econometrics*, 8, S153–S173.

Das S.R., Sundaram R.K., 1999, Of Smiles and Smirks: A Term Structure Perspective, *Journal of Financial and Quantitative Analysis*, 34, 211–239.

Derman E., Kani I., 1994, Riding on the Smile, *Risk*, 7, 32–39.

Derman E., Kani I., 1998, Stochastic Implied Trees: Arbitrage Pricing with Stochastic Term and Strike Structure of Volatility, *International Journal of Theoretical and Applied Finance*, 1, 61–110.

Detemple J., Osakwe C., 2000, The Valuation of Volatility Options, *European Finance Review*, 4, 21–50.

Duan J.C., 1995, The GARCH Option Pricing Model, *Mathematical Finance*, 5, 13–32.

Duan J.C., 1996, A Unified Theory of Option Pricing under Stochastic Volatility – from GARCH to Diffusion, Working Paper.

Duan J.C., 1997, Augmented GARCH(p,q) Process and its Diffusion Limit, *Journal of Econometrics*, 79, 97–127.

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

Duffie D., Kan R., 1996, A Yield-Factor Model of Interest Rates, *Mathematical Finance*, 6, 379–406.

Duffie D., Pan J., Singleton K., 2000, Transform Analysis and Asset Pricing for Affine Jump-Diffusions, *Econometrica*, 68, 1343–1376.

Duffie D., Singleton K., 1993, Simulated Moments Estimation of Markov Models of Asset Prices, *Econometrica*, 61, 929–952.

Dumas B., Fleming J., Whaley R.E., 1998, Implied Volatility Functions: Empirical Tests, *Journal of Finance*, 53, 2059–2106.

Dupire B., 1994, Pricing with a Smile, *Risk*, 7, 18–20.

Dupire B., 1997, Pricing and Hedging with Smiles, in: Dempster M.A.H., Pliska S.R. (editors), *Mathematics of Derivative Securities*, Cambridge University Press, Cambridge, pp. 103–111.

Durbin J., Koopman S.J., 1997, Monte Carlo Maximum Likelihood Estimation for Non-Gaussian State Space Models, *Biometrika*, 84, 669–684.

Durham G.B., Gallant A.R., 2002, Numerical Techniques for Maximum Likelihood Estimation of Continuous-Time Diffusion Processes, *Journal of Business and Economic Statistics*, 20, 297–316.

Elerian O., Chib S., Shephard N., 2001, Likelihood Inference for Discretely Observed Nonlinear Diffusions, *Econometrica*, 69, 959–993.

Engle R.F., Mustafa C., 1992, Implied ARCH Models from Options Prices, *Journal of Econometrics*, 52, 289–311.

Eraker B., 2001, MCMC Analysis of Diffusion Models with Application to Finance, *Journal of Business and Economic Statistics*, 19, 177–191.

Eraker B., Johannes M., Polson N., 2003, The Impact of Jumps in Volatility and Returns, *Journal of Finance*, 58, 1269–1300.

Fama E., 1965, The Behavior of Stock-Market Prices, *Journal of Business*, 38, 34–105.

Fan R., Gupta A., Ritchken P., 2002, Hedging in the Possible Presence of Unspanned Stochastic Volatility: Evidence from Swaption Markets, Working Paper.

Fermanian J.-D., Salanie B., 2004, A Nonparametric Simulated Maximum Likelihood Estimation Method, *Econometric Theory*, 20, 701–734.

Filipovic D., 2000, *Consistency Problems for HJM Interest Rate Models*, PhD Dissertation, Swiss Federal Institute of Technology.

Fong H.G., Vasicek O.A., 1991, Fixed-Income Volatility Management, *Journal of Portfolio Management*, 17, 41–46.

Gallant A.R., Long J.R., 1997, Estimating Stochastic Differential Equations Efficiently by Minimum Chi-Square, *Biometrika*, 84, 125–141.

Gallant A.R., Tauchen G.E., 1996, Which Moments to Match?, *Econometric Theory*, 12, 657–681.

Garcia R., Ghysels E., Renault E., 2004, The Econometrics of Option Pricing, CIRANO Working Paper 2004s–04.

Garcia R., Luger R., Renault E., 2002, Empirical Assessment of an Intertemporal

Option Pricing Model with Latent Variables, Working Paper.

Garman M.B., 1976, A General Theory of Asset Valuation under Diffusion State Processes, Working Paper No 50, Research Program in Finance, University of California, Berkeley.

Geman H., El Karoui N., Rochet J.-C., 1995, Changes of Numeraire, Changes of Probability Measure and Option Pricing, *Journal of Applied Probability*, 32, 443–458.

Geske R., 1979, The Valuation of Compound Options, *Journal of Financial Economics*, 7, 63–81.

Geweke J., 1989, Bayesian Inference in Econometric Models Using Monte Carlo Integration, *Econometrica*, 57, 1317–1339.

Glasserman P., Kou S., 2003, The Term Structure of Simple Forward Rates with Jump Risk, *Mathematical Finance*, 13, 383–410.

Goldstein R.S., 2000, The Term Structure of Interest Rates as a Random Field, *Review of Financial Studies*, 13, 365–384.

Gourieroux C., Monfort A., 1997, *Simulation-Based Econometric Methods*, Oxford University Press, Oxford.

Gourieroux C., Monfort A., Renault E., 1993, Indirect Inference, *Journal of Applied Econometrics*, 8, S85–S118.

Grunbichler A., Longstaff F.A., 1996, Valuing Futures and Options on Volatility, *Journal of Banking and Finance*, 20, 985–1001.

Hansen L.P., 1982, Large Sample Properties of Generalised Method of Moments Estimators, *Econometrica*, 50, 1029–1054.

Harvey A.C., Ruiz E., Shephard N.G., 1994, Multivariate Stochastic Variance Models, *Review of Economic Studies*, 61, 247–264.

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

Heidari M., Wu L., 2003, Are Interest Rate Derivatives Spanned by the Term Structure of Interest Rates?, *Journal of Fixed Income*, 13, 75–86.

Heston S.L., 1993, A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *Review of Financial Studies*, 6, 327–343.

Ho M.S., Perraudin W.R.M., Sorensen B.E., 1996, A Continuous-Time Arbitrage-Pricing Model with Stochastic Volatility and Jumps, *Journal of Business and*

Economic Statistics, 14, 31–43.

Ho T.S.Y., Lee S.B., 1986, Term Structure Movements and Pricing Interest Rate Contingent Claims, *Journal of Finance*, 41, 1011–1029.

Hobson D.G., Rogers L.C.G., 1998, Complete Models with Stochastic Volatility, *Mathematical Finance*, 8, 27–48.

Hull J., White A., 1987, The Pricing of Options on Assets with Stochastic Volatilities, *Journal of Finance*, 42, 281–300.

Ikeda N., Watanabe S., 1989, *Stochastic Differential Equations and Diffusion Processes*, Kodansha, Tokyo.

Jacquier E., Polson N.G., Rossi P.E., 1994, Bayesian Analysis of Stochastic Volatility Models, *Journal of Business and Economic Statistics*, 12, 371–417.

Jamshidian F., 1997, LIBOR and Swap Market Models and Measures, *Finance and Stochastics*, 1, 293–330.

Jarrow R., Li H., Zhao F., 2003, Interest Rate Caps "Smile" Too! But Can the LIBOR Market Models Capture It?, Working Paper.

Jeffrey A., Kristensen D., Linton O., Nguyen T., Phillips P., 2003, Nonparametric Estimation of a Multifactor Heath-Jarrow-Morton Model: An Integrated Approach, Working Paper.

Jiang G.J., Knight J.L., 2002, Estimation of Continuous-Time Processes via the Empirical Characteristic Function, *Journal of Business and Economic Statistics*, 20, 198–212.

Johnson H., Shanno D., 1987, Option Pricing when the Variance is Changing, *Journal of Financial and Quantitative Analysis*, 22, 143–151.

Jones C., 2003, The Dynamics of Stochastic Volatility: Evidence from Underlying and Options Markets, *Journal of Econometrics*, 116, 181–224.

Joshi M., Rebonato R., 2001, A Stochastic-Volatility, Displaced-Diffusion Extension of the LIBOR Market Model, Working Paper, Quantitative Research Centre of the Royal Bank of Scotland Group.

Kallsen J., Taqqu M.S., 1998, Option Pricing in ARCH-Type Models, *Mathematical Finance*, 8, 13–26.

Kendall M., 1987, *Kendall's Advanced Theory of Statistics, Volume 1: Distribution Theory*, 5th Edition, Griffin, London.

Kennedy D.P., 1994, The Term Structure of Interest Rates as a Gaussian Random Field, *Mathematical Finance*, 4, 247–258.

Kennedy D.P., 1997, Characterizing Gaussian Models of the Term Structure of Interest

Rates, *Mathematical Finance*, 7, 107–116.

Kerkhof J., Pelsser A., 2002, Observational Equivalence of Discrete String Models and Market Models, *Journal of Derivatives*, 10, 55–61.

Kimmel R., 2004, Modeling the Term Structure of Interest Rates: A New Approach, *Journal of Financial Economics*, 72, 143–183.

Lamoureux C., Lastrapes W., 1993, Forecasting Stock-Return Variance: Toward an Understanding of Stochastic Implied Volatilities, *Review of Financial Studies*, 6, 293–326.

Ledoit O., Santa-Clara P., Yan S., 2002, Relative Pricing of Options with Stochastic Volatility, Working Paper.

Liesenfeld R., Richard J.F., 2003, Univariate and Multivariate Stochastic Volatility Models: Estimation and Diagnostics, *Journal of Empirical Finance*, 10, 505–531.

Longstaff F.A., Santa-Clara P., Schwartz E.S., 2001a, The Relative Valuation of Caps and Swaptions: Theory and Empirical Evidence, *Journal of Finance*, 56, 2067–2109.

Longstaff F.A., Santa-Clara P., Schwartz E.S., 2001b, Throwing Away a Billion Dollars: the Cost of Suboptimal Exercise Strategies in the Swaptions Market, *Journal of Financial Economics*, 62, 39–66.

Longstaff F.A., Schwartz E.S., 1992, Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model, *Journal of Finance*, 47, 1259–1282.

Mandelbrot B., 1963, The Variation of Certain Speculative Prices, *Journal of Business*, 36, 394–419.

Melino A., Turnbull S.M., 1990, Pricing Foreign Currency Options with Stochastic Volatility, *Journal of Econometrics*, 45, 239–265.

Merton R., 1973, The Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science*, 4, 141–183.

Merville L.J., Pieptea D.R., 1989, Stock-Price Volatility, Mean-Reverting Diffusion, and Noise, *Journal of Financial Economics*, 24, 193–214.

Miltersen K.R., Sandmann K., Sondermann D., 1997, Closed Form Solutions for Term Structure Derivatives with Log-Normal Interest Rates, *Journal of Finance*, 52, 409–430.

Musiela M., Rutkowski M., 1997, Continuous-Time Term Structure Models: Forward Measure Approach, *Finance and Stochastics*, 1, 261–291.

Naik V., 1993, Option Valuation and Hedging Strategies with Jumps in the Volatility of Asset Returns, *Journal of Finance*, 48, 1969–1984.

Nandi S., 1996, Pricing and Hedging Index Options Under Stochastic Volatility: An Empirical Examination, Working Paper 96–9, Federal Reserve Bank of Atlanta.

Nandi S., 1998, How Important Is the Correlation between Returns and Volatility in a Stochastic Volatility Model? Empirical Evidence from Pricing and Hedging in the S&P 500 Index Options Market, *Journal of Banking and Finance*, 22, 589–610.

Nelson D.B., 1990, ARCH Models as Diffusion Approximations, *Journal of Econometrics*, 45, 7–38.

Pan J., 2002, The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study, *Journal of Financial Economics*, 63, 3–50.

Pearson N.D., Zhou A., 1999, A Nonparametric Analysis of the Forward Rate Volatilities, OFOR Paper Number 99–05, Office for Futures and Options Research of the University of Illinois.

Pedersen A.R., 1995, A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations, *Scandinavian Journal of Statistics*, 22, 55–71.

Press W.H., Teukolsky S.A., Vetterling W.T., Flannery B.P., 2002, *Numerical Recipes in C++*, Cambridge University Press, Cambridge.

Rao C.R., 1973, *Linear Statistical Inference and its Applications*, 2nd edition, Wiley, New York.

Rebonato R., 2003, Term-Structure Models: A Review, Working Paper, Quantitative Research Centre of the Royal Bank of Scotland Group.

Rebonato R., Cooper I., 1996, The Limitations of Simple Two-Factor Interest Rate Models, *Journal of Financial Engineering*, 5, 1–16.

Richard J.F., Zhang W., 2000, Accelerated Monte Carlo Integration: An Application to Dynamic Latent Variable Models, in: Mariano R., Schuermann T., Weeks M.J. (editors), *Simulation-Based Inference in Econometrics*, Cambridge University Press, Cambridge, pp. 47–70.

Romano M., Touzi N., 1997, Contingent Claims and Market Completeness in a Stochastic Volatility Model, *Mathematical Finance*, 7, 399–410.

Rubinstein M., 1983, Displaced Diffusion Option Pricing, *Journal of Finance*, 38, 213–217.

Rubinstein M., 1994, Implied Binomial Trees, *Journal of Finance*, 49, 771–818.

Santa-Clara P., Sornette D., 2001, The Dynamics of the Forward Rate Curve with Stochastic String Shocks, *Review of Financial Studies*, 14, 149–185.

Schobel R., Zhu J., 1999, Stochastic Volatility with an Ornstein-Uhlenbeck Process: an

Extension, *European Finance Review*, 3, 23–46.

Schonbucher P.J., 1999, A Market Model for Stochastic Implied Volatility, *Philosophical Transactions of the Royal Society A*, 357, 2071–2092.

Scott D.W., 1992, *Multivariate Density Estimation*, John Wiley, New York.

Scott L.O., 1987, Option Pricing when the Variance Changes Randomly: Theory, Estimation, and an Application, *Journal of Financial and Quantitative Analysis*, 22, 419–438.

Scott L.O., 1997, Pricing Stock Options in a Jump-Diffusion Model with Stochastic Volatility and Interest Rates: Applications of Fourier Inversion Methods, *Mathematical Finance*, 7, 413–424.

Shephard N., 1991, From Characteristic Function to Distribution Function: A Simple Framework for the Theory, *Econometric Theory*, 7, 519–529.

Silverman B.W., 1986, *Density Estimation*, Chapman and Hall, London.

Singleton K.J., 2001, Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function, *Journal of Econometrics*, 102, 111–141.

Skiadopoulos G., 2001, Volatility Smile Consistent Option Models: A Survey, *International Journal of Theoretical and Applied Finance*, 4, 403–437.

Stein E.M., Stein J.C., 1991, Stock Price Distributions with Stochastic Volatility: An Analytic Approach, *Review of Financial Studies*, 4, 727–752.

Vasicek O., 1977, An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics*, 5, 177–188.

Wiggins J.B., 1987, Option Values under Stochastic Volatility, Theory and Empirical Estimates, *Journal of Financial Economics*, 19, 351–372.

Appendices

A. The Ito-Venttsel formula

Let $W_i(t)$, $i = 0, 1, \ldots, n$, be independent standard Brownian motions, and $f(t, u)$ a stochastic process with the following dynamics when u is not stochastic:

$$df(t,u) = \alpha(t,u)\,dt + \sum_{i=0}^{n} \beta_i(t,u)\,dW_i(t)$$

If $v(t)$ is a stochastic process with dynamics

$$dv(t) = \gamma(t,v)\,dt + \sum_{i=0}^{n} \delta_i(t,v)\,dW_i(t)$$

then $f[t, v(t)]$ satisfies the following stochastic partial differential equation:

$$\begin{aligned} df[t,v(t)] &= \alpha[t,v(t)]\,dt + \sum_{i=0}^{n} \beta_i[t,v(t)]\,dW_i(t) + \frac{\partial f}{\partial v} dv(t) \\ &\quad + \frac{1}{2}\frac{\partial^2 f}{\partial v^2} \sum_{i=0}^{n} \delta_i^2(t,v)\,dt + \sum_{i=0}^{n} \frac{\partial \beta_i}{\partial v} \delta_i(t,v)\,dt \end{aligned} \tag{A.1}$$

(See Brace *et al.* (2001) and references therein.)

B. Proof of lemma 2.3

Equation (2.36) is obtained by a simple extension of Shephard's (1991) proof for formula (2.33). Let $G(y) = \text{sign}(y)$, $y \in [-h, h]$. The Fourier transform of $G(y)$ is

$$\Phi(\varphi; h) = \int_{-h}^{h} G(y)\, e^{i\varphi y} dy = \frac{2[\cos(h\varphi) - 1]}{i\varphi}$$

The convolution of $G(y)$ with the product $p(y)\,e^{\alpha y}$, where $p(y)$ is the continuous probability density function of y, is equal to:

$$\Psi(K; h) = \int_{-h}^{h} p(y)\, e^{ay} G(K - y)\, dy = \int_{-h}^{K} p(y)\, e^{\alpha y} dy - \int_{K}^{h} p(y)\, e^{\alpha y} dy$$

where $h > K$ and $h > 0$. The characteristic function of y is:

$$f(\varphi) = \int_{-\infty}^{\infty} e^{i\varphi y} p(y)\, dy$$

The Fourier transform of $\Psi(K; h)$ is equal to:

$$\frac{2[\cos(h\varphi) - 1]}{i\varphi} \int_{-\infty}^{\infty} e^{i\varphi y + \alpha y} p(y)\, dy = \frac{2[\cos(h\varphi) - 1]}{i\varphi} f(\varphi - i\alpha)$$

where we use the fact that the Fourier transform of a convolution of two functions is equal to the product of the Fourier transforms of the two functions. By inverting the Fourier transform of $\Psi(K; h)$ we obtain:

$$\begin{aligned}
\lim_{h\to\infty} \Psi(K; h) &= E(e^{\alpha y}) - 2E(e^{\alpha y} I_{y\geq K}) \\
&= \frac{2}{2\pi} \lim_{h\to\infty} \int_{-\infty}^{\infty} e^{-i\varphi K} \frac{[\cos(h\varphi) - 1]}{i\varphi} f(\varphi - i\alpha)\, d\varphi \\
&= \frac{2}{\pi} \int_0^{\infty} \lim_{h\to\infty} [\cos(h\varphi) - 1] \operatorname{Re}\left[\frac{f(\varphi - i\alpha) e^{-i\varphi K}}{i\varphi}\right] d\varphi \\
&= -\frac{2}{\pi} \int_0^{\infty} \operatorname{Re}\left[\frac{f(\varphi - i\alpha) e^{-i\varphi K}}{i\varphi}\right] d\varphi
\end{aligned}$$

Combining this with the relationship $E(e^{\alpha y}) = f(-i\alpha)$, we obtain the required result. □

C. Proof of proposition 3.1

(Goldstein, 2000) Let $U(t, s) = -\int_t^s f(t, x)\, dx$. Then $P(t, s) = e^{U(t,s)}$. The dynamics of $U(t, s)$ are:

$$d_t U(t, s) = r(t)\, dt - \int_t^s \alpha(t, y)\, dy dt - \int_t^s \sigma(t, y)\, d_t Z(t, y)\, dy$$

Using Ito's lemma,

$$\begin{aligned}
\frac{d_t P(t, s)}{P(t, s)} &= d_t U(t, s) + \frac{1}{2} d_t \langle U(t, s) \rangle \\
&= r(t)\, dt - \int_t^s \alpha(t, y)\, dy dt - \int_t^s \sigma(t, y)\, d_t Z(t, y)\, dy \\
&\quad + \frac{1}{2} \int_t^s \int_t^s \sigma(t, y)\, \sigma(t, x)\, c(x, y)\, dy dx dt
\end{aligned}$$

Under the risk-neutral measure the drift of bond prices must be equal to the spot rate:

$$\frac{d_t P(t, s)}{P(t, s)} = r(t)\, dt - \int_t^s \sigma(t, x)\, d_t Z(t, x)\, dx$$

Therefore, the following condition must hold for no arbitrage opportunities to exist:

$$\int_t^s \alpha(t, y)\, dy = \frac{1}{2} \int_t^s \int_t^s \sigma(t, y)\, \sigma(t, x)\, c(x, y)\, dy dx$$

By differentiating the above equation with respect to s we obtain the required result.

□

D. Proof of proposition 3.2

(Goldstein, 2000) Suppose that $V(t)$ represents the value at time t of a portfolio consisting of η shares of the bank account $B(t)$ and θ_u shares of the zero-coupon risk-free bond maturing at time u, $\forall\ u > t$:

$$V(t) = \eta B(t) + \int_t^{\infty} \theta_u P(t,u)\, du$$

Hence,

$$dV(t) = \eta dB(t) + \int_t^{\infty} \theta_u d_t P(t,u)\, du$$

Under the risk-neutral measure $\mathbb{Q}$, $\frac{V(t)}{B(t)}$ is a martingale. Under the T-forward measure $\mathbb{Q}^T$, $V^T(t) = \frac{V(t)}{P(t,T)}$ is a martingale. Using Ito's lemma we obtain the dynamics of $V^T(t)$:

$$\begin{aligned} dV^T(t) &= \left[\frac{\eta B(t)}{P(t,T)} \int_t^T \sigma(t,x)\, dx + \int_t^{\infty} \theta_u \frac{P(t,u)}{P(t,T)} \int_u^T \sigma(t,x)\, du dx\right] \\ &\times \left[d_t Z(t,x) + \int_t^T \sigma(t,y)\, c(x,y)\, dy dt\right] \end{aligned}$$

The above equation implies that the T-forward measure is given by equation (3.10).

□

E. The forward-rate drift under the risk-neutral measure in section 4.2

(Santa-Clara and Sornette, 2001) The following relationship holds between bond prices and forward rates:

$$\ln P(t,u) = -\int_0^{u-t} f(t,x)\, dx \tag{E.1}$$

From the above equation it follows that:

$$d_t \ln P(t,u) = f(t,u-t)\, dt - \int_0^{u-t} d_t f(t,x)\, dx$$

Using Ito's lemma we obtain the dynamics of bond prices:

$$\begin{aligned} \frac{d_t P(t,u)}{P(t,u)} &= \left[f(t,u-t) + \frac{1}{2}\int_0^{u-t}\int_0^{u-t} \sigma(t,x)\,\sigma(t,y)\, c(x,y)\, dx dy\right] dt \\ &- \int_0^{u-t} \alpha(t,x)\, dx dt - \int_0^{u-t} \sigma(t,x)\, d_t Z_f(t,x)\, dx \end{aligned}$$

Under the risk-neutral measure the drift of bond prices must be equal to the spot rate. Therefore,

$$\int_0^{u-t} \alpha(t,x)\,dx = f(t,u-t) + \frac{1}{2}\int_0^{u-t}\int_0^{u-t} \sigma(t,x)\,\sigma(t,y)\,c(x,y)\,dxdy - r(t)$$

Let $s = u - t$. Taking the derivative with respect to u in the above equation we obtain:

$$\alpha(t,s) = \frac{\partial f(t,s)}{\partial s} + \sigma(t,s)\int_0^{s} \sigma(t,x)\,c(s,x)\,dx \qquad \square \tag{E.2}$$

F. Graphs of exponential functions with estimated parameters in models I, II, and III

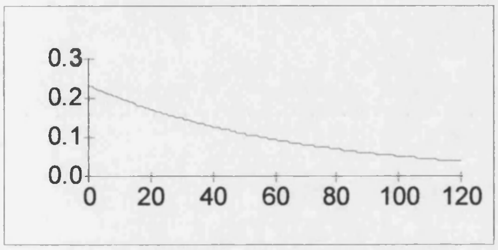

Figure 1. Model I UK *g(s)* function

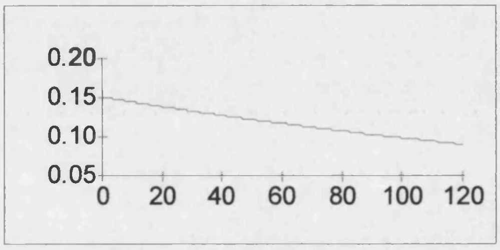

Figure 2. Model I US *g(s)* function

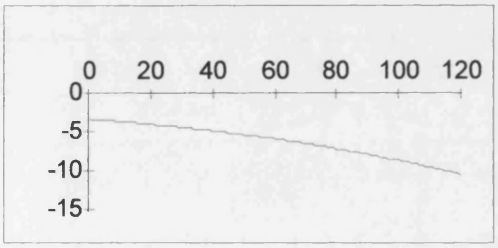

Figure 3. Model I UK *R(s)* function

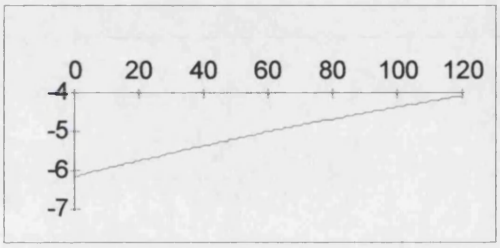

Figure 4. Model I US *R(s)* function

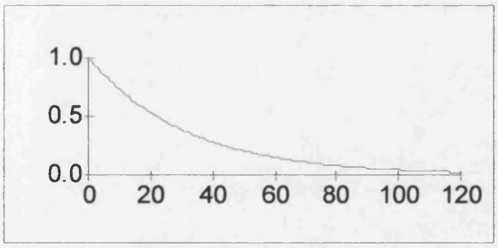

Figure 5. Model I UK $c(s_1\text{-}s_2)$ function

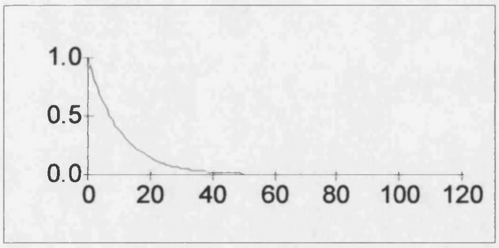

Figure 6. Model I US $c(s_1\text{-}s_2)$ function

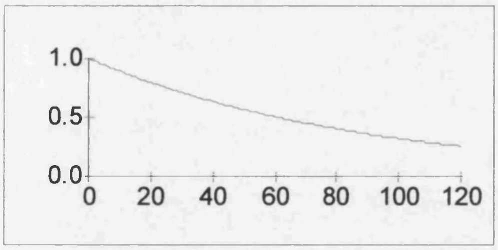

Figure 7. Model II UK *g(s)* function

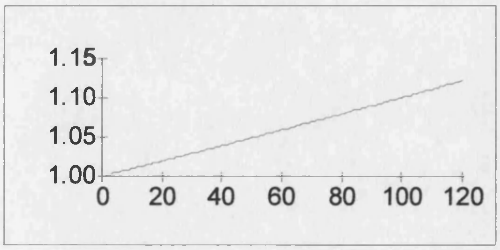

Figure 8. Model II US *g(s)* function

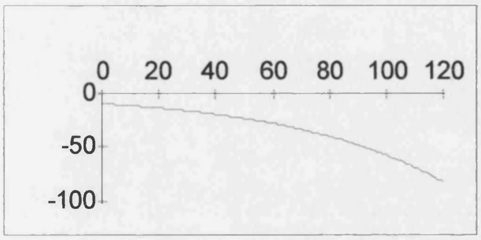

Figure 9. Model II UK *R(s)* function

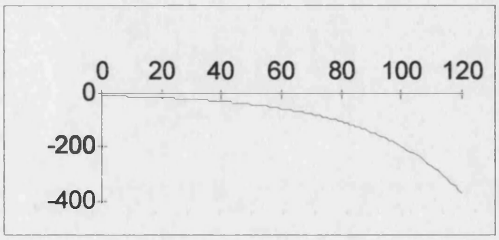

Figure 10. Model II US *R(s)* function

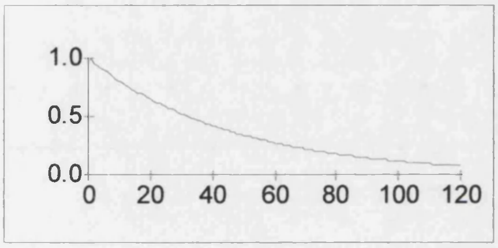

Figure 11. Model II UK $c(s_1\text{-}s_2)$ function

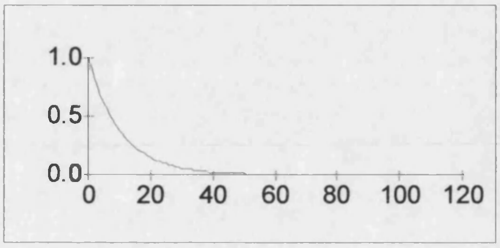

Figure 12. Model II US $c(s_1\text{-}s_2)$ function

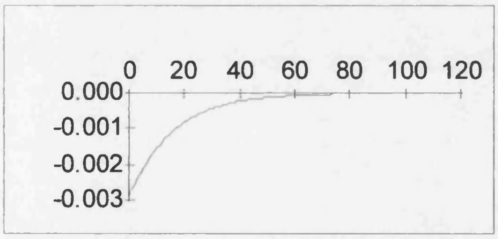

Figure 13. Model III UK *R(s)* function

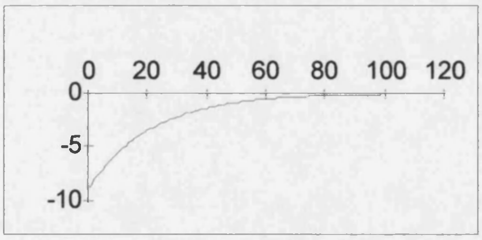

Figure 14. Model III US *R(s)* function

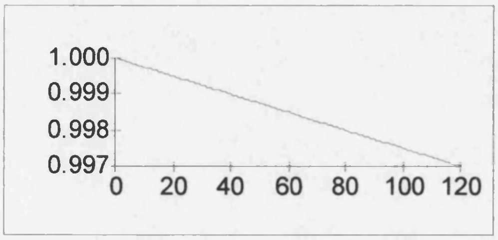

Figure 15. Model III UK $c(s_1\text{-}s_2)$ function

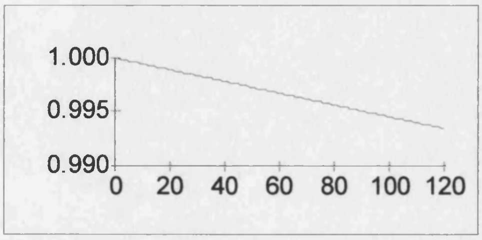

Figure 16. Model III US $c(s_1\text{-}s_2)$ function

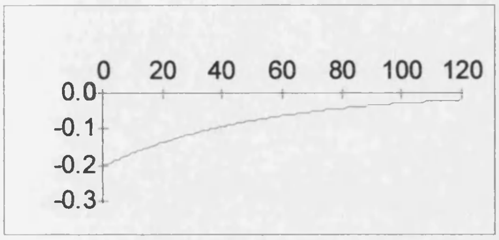

Figure 17. Model III UK $\alpha(s)$ function

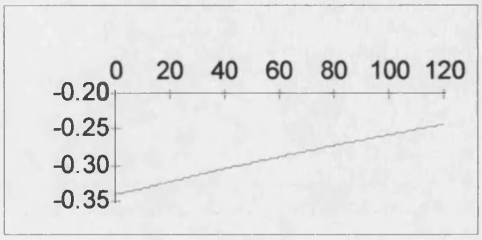

Figure 18. Model III US $\alpha(s)$ function

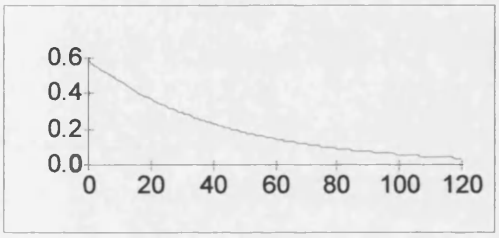

Figure 19. Model III UK $\xi(s)$ function

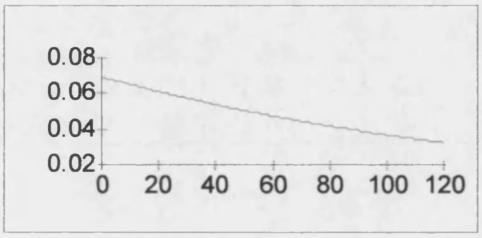

Figure 20. Model III US $\xi(s)$ function

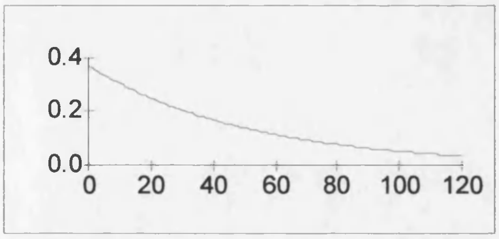

Figure 21. Model III UK $\phi(s)$ function

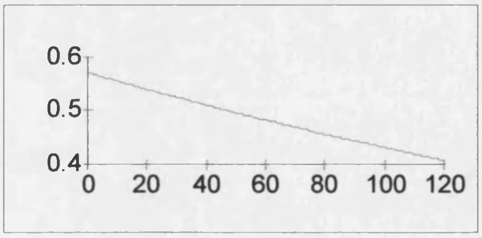

Figure 22. Model III US $\phi(s)$ function

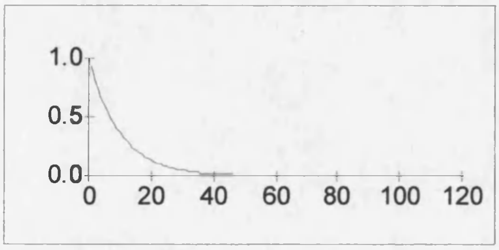

Figure 23. Model III UK $\rho(s_1-s_2)$ function

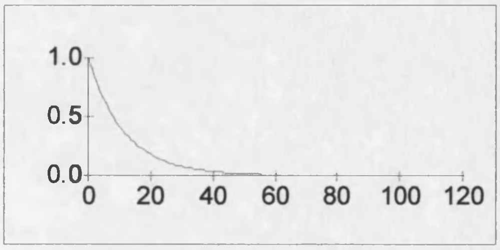

Figure 24. Model III US $\rho(s_1-s_2)$ function

G. Monte Carlo study results

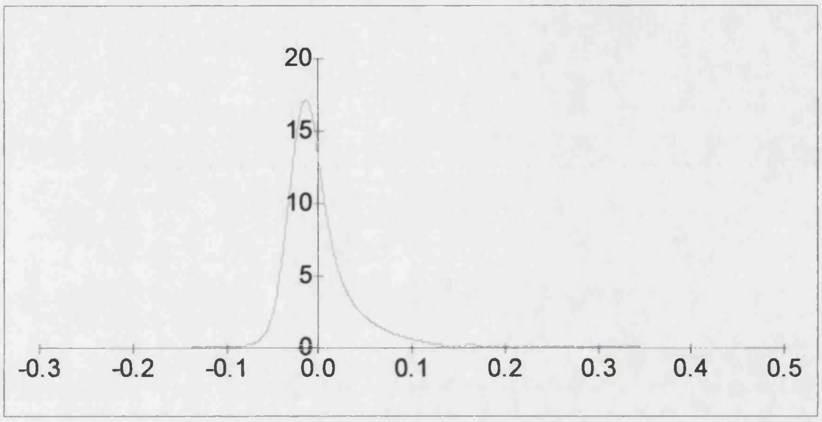

Figure 25. Distribution of estimated values for φ_0

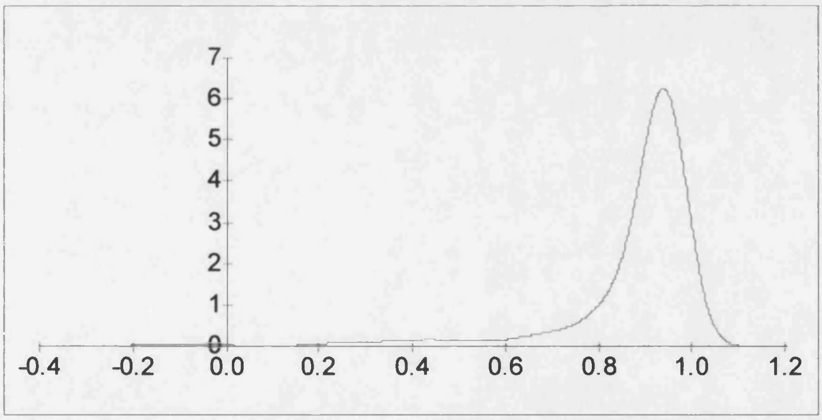

Figure 26. Distribution of estimated values for φ_1

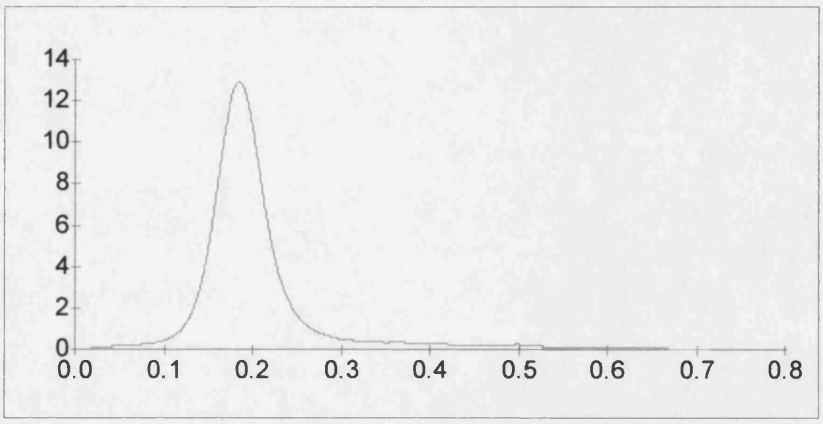

Figure 27. Distribution of estimated values for σ