Pricing Long-Dated Equity Derivatives under Stochastic Interest Rates

Navin Ranasinghe

Submitted in total fulfillment of the requirements of the degree of Doctor of Philosophy

December, 2016
Centre for Actuarial Studies
Department of Economics
The University of Melbourne

Produced on archival quality paper
Abstract

A key requirement of any equity hybrid derivatives pricing model is the ability to rapidly and accurately calibrate to vanilla option prices. However, existing methodologies are often reliant on costly numerical procedures or approximations that may not be suitable when dealing with long-term expiries. Therefore, in this thesis, we introduce new techniques for calibrating equity models under correlated stochastic interest rates, which do not suffer from these limitations. We also present a number of empirical examples to highlight the potential impact of interest rate stochasticity on long-dated derivatives.

In chapter 3, we begin by introducing a class of equity hybrid models that is capable of producing an implied volatility smile. This is achieved by equating the stock price divided by the bank account to a chosen function of a driving Gaussian process. The resulting processes for the stock price, short-rate and bank account can be exactly simulated over arbitrary time steps because they follow a straightforward transformation of the joint normal distribution. Furthermore, vanilla option prices are available as a one dimensional integral, meaning that these models can be efficiently calibrated.

However, under our approach, the function linking the stock price to the driving Gaussian process is not allowed to vary arbitrarily with time, and must instead be chosen to satisfy a particular no arbitrage condition. This restriction means that there is only a single time-dependent parameter, the volatility of the driving Gaussian process, and it may struggle to match vanilla option prices across multiple expiries. We address this issue in chapter 4 by showing how to construct mixture models, under non-deterministic interest rates, which use the models developed in chapter 3 as the underlying components. These mixture models allow for an arbitrary number of time-dependent parameters, and may therefore be accurately calibrated to the entire implied volatility surface.

Building on this, in chapter 5, we extend our mixture-based approach to include stochastic volatility, in addition to local volatility and stochastic interest rates. This requires deriving the joint characteristic function of a suitable class of component models, and then utilizing the multidimensional fractional FFT. Compared to those previously discussed, the resulting model allows for more realistic volatility dynamics, which is helpful when pricing certain exotic derivatives, such as forward start options and ratchet options.

On the other hand, when dealing with volatility derivatives, it is sometimes possible to write their price directly in terms of the prices of vanilla options across all strikes and expiries, without
adopting a specific parametric model for the stock. The main benefit of this approach is that it will be consistent with any model satisfying the underlying assumptions, and will not depend on how the model is parametrized or calibrated. However, existing results either assume deterministic interest rates or do not apply to any volatility derivatives other than the standard variance swap. Thus, in chapter 6, we extend the non-parametric pricing of general volatility derivatives to the case of stochastic interest rates, given certain independence and continuity assumptions.
Declaration

This is to certify that:

1. *the thesis comprises only my original work towards the PhD except where indicated in the preface*;

2. *due acknowledgement has been made in the text to all other material used*;

3. *the thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.*

Signed,

Navin Ranasinghe
Preface

This thesis was produced under the supervision of Professor Mark Joshi at the Centre for Actuarial Studies, The University of Melbourne. Chapters 3 to 6 present its original contributions, except as stated otherwise in the text.

The research and writing of chapter 3 was done by Navin Ranasinghe, with supervision, proofreading and editing by Mark Joshi.

Chapter 4 is based on the paper “Local Volatility under Stochastic Interest Rates using Mixture Models”, which was co-authored by Mark Joshi. The research and writing was done by Navin Ranasinghe, with supervision, proofreading and editing by Mark Joshi.

Chapter 5 is based on the paper “Local and Stochastic Volatility under Stochastic Interest Rates using Mixture Models and the Multidimensional Fractional FFT”, which was co-authored by Mark Joshi. The research and writing was done by Navin Ranasinghe, with supervision, proofreading and editing by Mark Joshi.

Chapter 6 is based on the paper “Non-Parametric Pricing of Long-Dated Volatility Derivatives under Stochastic Interest Rates”, which was co-authored by Mark Joshi, and published in Quantitative Finance. The research and writing was done by Navin Ranasinghe, with supervision, proofreading and editing by Mark Joshi.

None of the work towards this thesis has been submitted for any other qualifications, nor was it carried out prior to enrolment in the degree. No specific grants from funding agencies in the public, commercial, or not-for-profit sectors were received for this research.
Acknowledgements

I am deeply indebted to Professor Mark Joshi for giving me the opportunity to complete my PhD under his supervision. His guidance regarding all parts of my research, from the fundamental concepts right down to the smallest details, is what made this thesis possible. I am especially grateful for his thorough and insightful comments on numerous drafts, and his careful questioning that helped turn vague ideas into concrete results. I simply cannot imagine a better mentor for an aspiring financial mathematician.

My sincere thanks go to all the staff in the Centre for Actuarial Studies at the University of Melbourne for the education and assistance they provided me throughout my graduate and undergraduate studies. I also wish to express my heartfelt gratitude to my fellow PhD students for always providing a friendly and intellectually stimulating environment.

Finally, I would like to thank my parents for their continuous and unparalleled support, not just in my studies, but in all aspects of my life. I would never have taken the risk to pursue this degree without their unwavering belief in my ability and constant encouragement to fulfil my potential. For this, and so many other things, I will forever be grateful to them.
Contents

1 Introduction ................................................................. 1
   1.1 Motivation .......................................................... 1
   1.2 The Black-Scholes Model ........................................... 2
   1.3 Local Volatility ....................................................... 3
   1.4 Stochastic Volatility ............................................... 4
   1.5 Mixture Models ..................................................... 6
   1.6 Stochastic Interest Rates ........................................... 8
   1.7 Non-Parametric Pricing of Volatility Derivatives ................. 10
   1.8 Outline of the Monograph .......................................... 11

2 Review of Equity Derivatives Pricing under Stochastic Interest Rates 13
   2.1 Local Volatility under Stochastic Interest Rates ................... 14
   2.2 Stochastic Volatility under Stochastic Interest Rates ............. 16
   2.3 Combined Local and Stochastic Volatility ......................... 17
   2.4 Volatility Derivatives under Stochastic Interest Rates .......... 19
   2.5 Conclusion ......................................................... 21

3 Parametric Local Volatility Models under Stochastic Interest Rates 23
   3.1 Introduction ......................................................... 23
   3.2 Assumptions and Main Results ..................................... 24
   3.3 Example Models ..................................................... 30
   3.4 Empirical Results .................................................. 36
   3.5 Conclusion ........................................................ 39
   3.6 Proofs ............................................................. 39

4 Local Volatility under Stochastic Interest Rates Using Mixture models 43
   4.1 Introduction ........................................................ 43
   4.2 Approximate Local Volatility Using Mixture Models .............. 44
   4.3 Multivariate Local Volatility Using Mixture Models ............... 47
   4.4 Example Mixture Model ....................................... 51
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.5 Empirical Results</td>
<td>53</td>
</tr>
<tr>
<td>4.6 Conclusion</td>
<td>59</td>
</tr>
<tr>
<td>5 Local and Stochastic Volatility under Stochastic Interest Rates Using Mixture Models</td>
<td>63</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>63</td>
</tr>
<tr>
<td>5.2 Mixtures of Stochastic Volatility Models under Stochastic Interest Rates</td>
<td>66</td>
</tr>
<tr>
<td>5.3 Empirical Results</td>
<td>73</td>
</tr>
<tr>
<td>5.4 Conclusion</td>
<td>75</td>
</tr>
<tr>
<td>5.A Proofs</td>
<td>76</td>
</tr>
<tr>
<td>6 Non-Parametric Pricing of Volatility Derivatives under Stochastic Interest Rates</td>
<td>81</td>
</tr>
<tr>
<td>6.1 Introduction</td>
<td>81</td>
</tr>
<tr>
<td>6.2 Notation and Assumptions</td>
<td>84</td>
</tr>
<tr>
<td>6.3 Exponential Variance Contracts</td>
<td>86</td>
</tr>
<tr>
<td>6.4 Correlation Neutrality</td>
<td>88</td>
</tr>
<tr>
<td>6.5 Other Variance Contracts</td>
<td>91</td>
</tr>
<tr>
<td>6.5.1 Power Payoffs</td>
<td>91</td>
</tr>
<tr>
<td>6.5.2 Payoffs with Exponentially Decaying Transforms</td>
<td>94</td>
</tr>
<tr>
<td>6.5.3 Other Payoff Functions</td>
<td>95</td>
</tr>
<tr>
<td>6.6 Unbounded Quadratic Variation</td>
<td>96</td>
</tr>
<tr>
<td>6.7 Mixture of Normals Method</td>
<td>97</td>
</tr>
<tr>
<td>6.7.1 Comparison to Existing Fitting Procedures</td>
<td>100</td>
</tr>
<tr>
<td>6.8 Empirical Results</td>
<td>101</td>
</tr>
<tr>
<td>6.9 Conclusion</td>
<td>102</td>
</tr>
<tr>
<td>6.A Proofs</td>
<td>103</td>
</tr>
<tr>
<td>7 Summary and Conclusion</td>
<td>107</td>
</tr>
<tr>
<td>References</td>
<td>111</td>
</tr>
</tbody>
</table>
List of Figures

3.1 Shifted exponential model calibration results ........................................... 37
3.2 Hyperbolic sine model calibration results .................................................. 37
3.3 Modified exponential model calibration results .......................................... 38

4.1 BSHW mixture model with different drifts: calibration for $\rho = 0.4$ ............... 55
4.2 Shifted exponential mixture model with identical drifts: calibration for $\rho = 0.4$ . 55
4.3 Modified exponential mixture model with identical drifts: calibration for $\rho = 0.4$ . 55

5.1 SZHW mixture model calibration results for $\rho_{1,2} = 0.4$ .......................... 73
# List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Hull-White model calibration results: cap implied volatilities</td>
<td>37</td>
</tr>
<tr>
<td>4.1</td>
<td>Monte Carlo prices computed using the approximate local volatility function given in proposition 4.2.1</td>
<td>57</td>
</tr>
<tr>
<td>4.2</td>
<td>Monte Carlo prices, computed using the approximate local volatility function given in proposition 4.2.1, after fitting to adjusted market prices</td>
<td>58</td>
</tr>
<tr>
<td>4.3</td>
<td>Monte Carlo prices for $\rho = 0.4$, computed using the approximate local volatility function given in proposition 4.2.1, after fitting to market prices adjusted for a second time</td>
<td>59</td>
</tr>
<tr>
<td>4.4</td>
<td>Prices of at-the-money up-and-out call options valued at 28 April 2015 and expiring on 18 December 2020, computed using the methodology of section 4.4</td>
<td>60</td>
</tr>
<tr>
<td>5.1</td>
<td>Prices of at-the-money up-and-out call options, valued at 28 April 2015 and expiring on 18 December 2020, computed using Monte Carlo simulation of the model given in theorem 5.2.3</td>
<td>74</td>
</tr>
<tr>
<td>5.2</td>
<td>Prices of out-of-the-money vanilla options for $\rho_{1,2} = 0.4$, valued at 28 April 2015 and expiring on 18 December 2020, computed using Monte Carlo simulation of the model given in theorem 5.2.3</td>
<td>75</td>
</tr>
<tr>
<td>6.1</td>
<td>Fair strikes and prices of volatility derivatives valued at 21 October 2014 and expiring on 15 December 2023</td>
<td>102</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Motivation

Long-dated equity derivatives are frequently used by life insurers, fund managers, and other financial institutions to manage risks and provide tailored investment products to their clients. Although the effects of stochastic interest rates are often ignored when dealing with short-term contracts, they become increasingly significant as the term increases. Furthermore, it is necessary to jointly model stock prices and interest rates when pricing hybrid derivatives that explicitly depend on both of these quantities. Thus, our aim in this monograph is to extend existing derivatives pricing techniques, specifically local volatility, stochastic volatility, and model free pricing, to allow for non-deterministic interest rates.

Although the need for such extensions when pricing hybrid derivatives is obvious, their importance when pricing long-dated path-dependent derivatives, which do not directly depend on interest rates, is less well understood. For example, when replicating volatility derivatives, it is common practice to treat interest rates as deterministic, based on the assumption that the volatility of bond prices is not significant compared to the volatility of equities. Although this assumption is fine for short expiries, it is not at all safe when dealing with expiries many years into the future. In fact, we will give various empirical examples that highlight the potential impact of interest rate stochasticity on long-dated equity derivatives.

When developing a derivatives pricing methodology, two of the most important requirements are to model the stochastic nature of the underlying variables in a believable way, and to accurately reproduce the observed market prices of liquid instruments. However, when trying to achieve these goals, it is often necessary to resort to approximate techniques or computationally expensive algo-
rithms during calibration. Conversely, we will develop models that can be rapidly and accurately calibrated to market data, while maintaining the complexity required to provide a sufficiently realistic representation of the dynamics of interest rates and stock prices.

The starting point for our discussion is the ubiquitous options pricing model of Black and Scholes (1973). The two key assumptions of this model that we wish to relax are that interest rates are deterministic and that volatility is deterministic. In the literature, two common extensions dealing with this latter assumption are local volatility and stochastic volatility. The first of these allows volatility to be a function of the stock price as well as time, while the second allows volatility to follow its own stochastic process. In chapters 3, 4 and 5, our overall goal is to combine these extensions with stochastic interest rates, whereas in chapter 6 we pursue an alternative non-parametric approach. However, before going into further details, we summarise the Black-Scholes model in the following section.

1.2 The Black-Scholes Model

The Black-Scholes model is the basis of much work in mathematical finance. The goal of this model is to determine the price of an option, \( V \), that pays the owner \( V(S_T, T) \) dollars at expiry time \( T \). We begin with the assumption that the stock price, \( S_t \), follows geometric Brownian motion in the real-world measure, i.e.

\[
dS_t = \mu S_t dt + \sigma S_t dW_t,
\]

where \( \mu, \sigma \in \mathbb{R} \) are the drift and volatility of the stock price, and \( W_t \) is a standard Brownian motion adapted to the filtration \( \mathcal{F}_t \). By constructing a risk-free portfolio containing the option \( V \) and a variable number of stocks, and then equating the drift of this portfolio to the risk-free rate, \( r \), it is possible to derive the famous Black-Scholes equation,

\[
\frac{\partial V}{\partial t}(S, t) + rS \frac{\partial V}{\partial S}(S, t) + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2}(S, t) - rV(S, t) = 0. \tag{1.2.1}
\]

This equation can then be solved using the appropriate boundary conditions to yield the price of the option at time zero. As an alternative to this PDE based approach, it can be shown that there exists an equivalent “risk-neutral” probability measure under which the value of the option divided by the value of the bank account is a martingale. In this measure the drift of a non-dividend paying stock must equal the risk-free rate, i.e. \( dS_t = rS_t dt + \sigma S_t dW_t \), and the price of our option is

\[
V(S_t, t) = \mathbb{E}\left( e^{-rt}V(S_T, T) \mid \mathcal{F}_t \right), \tag{1.2.2}
\]

where the expectation is taken in the risk-neutral measure rather than the real-world measure. Throughout this monograph we adopt this martingale pricing approach, instead of the PDE based approach. In the case of a call option with expiry \( T \) and strike \( K \), which has pay-off \( C(S_T, T) = \)


(1.3. Local Volatility)

\((S_T - K)^+\), equation (1.2.2) yields the well known Black-Scholes formula

\[
C(S_t, t) = S_t N(d_1) - Ke^{-r(T-t)} N(d_1),
\]

\[
d_1 = \frac{1}{\sigma \sqrt{T-t}} \left( \log \left( \frac{S_t}{K} \right) + \left( r + \frac{\sigma^2}{2} \right) (T-t) \right),
\]

\[
d_2 = d_1 - \sigma \sqrt{T-t},
\]

where \(N(x)\) is the standard normal cumulative distribution function, and \((x)^+ := \max(x, 0)\). For a more detailed introduction the topic of derivatives pricing, and it's mathematical foundations, we refer the reader to Baxter and Rennie (1996), Björk (2009), Joshi (2003), or Wilmott, Howison, and Dewynne (1995).

The Black-Scholes model is so entrenched in derivatives pricing that the values of call options are often quoted in terms of their “implied volatility”, which is the value \(\sigma\) that, when entered into formula (1.2.3), yields the market price of the option. If the model were true, we would expect this implied volatility to be constant, and therefore independent of both \(T\) and \(K\). However, in the real world, we find that implied volatilities vary with both of these variables. The dependence on \(T\) can easily be accounted for by extending the model to a time dependent risk-free rate, \(r_t\), and volatility, \(\sigma_t\). The only changes to the Black-Scholes formula necessary are to make the substitutions

\[
r = \frac{1}{T-t} \int_t^T r_u du, \quad \sigma^2 = \frac{1}{T-t} \int_t^T \sigma_u^2 du.
\]

On the other hand, explaining the dependence of implied volatility on the strike is more difficult. One way to handle this phenomenon, which is known as the “implied volatility smile”, is to allow the stock price to follow a more general process than geometric Brownian motion. For example, one may allow the volatility, \(\sigma\), to depend on both the current stock price and time, which leads us to our next topic, the local volatility model.

### 1.3 Local Volatility

The local volatility model, introduced for continuous time by Dupire (1997), and for discrete time by Derman and Kani (1998), provides an effective way to account for the implied volatility smile. Compared to the Black-Scholes model, the key difference is that we replace the parameter \(\sigma_t\) with the “local volatility function” \(\sigma(S_t, t)\), i.e.

\[
dS_t = r_t S_t dt + \sigma(S_t, t) S_t dW_t
\]
in the risk-neutral measure. It can be shown that, given the complete surface of call prices, $C$, for all strikes, $K$, and expiries, $T$, the squared local volatility (which we call the local variance) is

$$\sigma^2(K, T) = \frac{\partial C}{\partial T} + rK \frac{\partial C}{\partial K} \frac{1}{2} K^2 \frac{\partial^2 C}{\partial K^2}. \quad (1.3.1)$$

In reality, call option prices are only available for a finite set of strikes and expiries in the market. Thus, before implementing this formula, it is first necessary to interpolate between observed call prices. However, as noted by Gatheral (2006), this interpolation needs to be done carefully so that the resulting surface is arbitrage free.

In addition to reproducing the observed market prices of vanilla options, the local volatility model has the convenient feature that there is only one source of randomness. This results in a complete market in which hedging options only requires the dynamic trading of shares. In general, adding additional sources of risk, as is done in stochastic volatility models, leads to hedging strategies that require the continuous trading of options as well as shares.

Nevertheless, the reliance on a single stochastic factor also leads to some undesirable properties. For example, the evolution of the implied volatility surface through time depends only on the movement of the stock price. This conflicts with the real world observation that this surface can change level or shape independently of changes in the stock price. Consequently, the local volatility model may significantly misprice options that depend directly on the dynamics of implied volatility, such as forward start options and ratchet options.

More generally, as observed by Dumas, Fleming, and Whaley (1998), the assumption that volatility is a deterministic function of the stock price is unrealistic and inconsistent with empirical evidence. Instead, if we are to have any hope of producing realistic dynamics for both stock prices and implied volatilities, we need to incorporate an additional stochastic factor into volatility.

### 1.4 Stochastic Volatility

As an alternative to the local volatility approach described above, we can instead let the volatility of the stock price follow its own stochastic process. In other words, we let

$$dS_t = r_t S_t dt + \eta_t S_t dW_{1,t},$$

where $\eta_t$ is itself stochastic. The goal of much research into stochastic volatility models is to identify specifications for $\eta_t$ which are realistic, can be calibrated rapidly, and produce implied volatility surfaces that match what is observed in the market. One of the most popular models in the literature is that of Heston (1993), under which variance, $v_t := \eta_t^2$, follows a mean-reverting square-root process, i.e.

$$dv_t = \kappa (\bar{v} - v_t) dt + \gamma \sqrt{v_t} dW_{2,t}, \quad dW_{1,t} dW_{2,t} = \rho dt,$$
where $\kappa$ is the rate of mean-reversion, $\bar{v}$ is the long-run mean, and $\gamma$ controls the volatility of volatility. Importantly, the model allows non-zero correlation between the driving Brownian motions, $W_{1,t}$ and $W_{2,t}$. As observed by Black (1975), downwards shocks to the stock price often coincide with upwards shocks to volatility, and vice versa, meaning that the aforementioned correlation is typically quite negative.

Observe that $\eta_t$ is both positive and mean-reverting, which is what we would expect of a realistic volatility process. The key result that makes this model tractable is that the characteristic function of the log stock price is known analytically. The original approach of Heston (1993) prices call options by decomposing them into a linear combination of two probabilities that can be computed by inverting this characteristic function.

A more recent approach, due to Carr and Madan (1999), is to work with the Fourier-Laplace transform of call option prices with respect to log-strike. This can be written in terms of the characteristic function of the log stock, and can be inverted numerically to recover call option prices. Conveniently, we can use the fast Fourier transform (FFT) to simultaneously compute prices for many different strikes, which makes calibration to a large number of strikes much faster.

This Fourier-Laplace transform based pricing procedure can be applied to any model for which the characteristic function of the log stock price is known. Consequently, much work has gone into characterizing the class of such models. For example, Duffie, Pan, and Singleton (2000) show that, for any jump diffusion model whose drift and instantaneous covariance matrix are an affine function of the state variables, the characteristic function can be derived by solving a certain set of coupled ordinary differential equations.

Although such models can be extended by adding more state variables (leading to multi-factor volatility models), their capacity to fit the market implied volatility surface is ultimately limited compared to the local volatility approach. When calibrating a stochastic volatility model we seek to minimize some measure of distance (e.g. the squared difference) between model and market call prices using a limited set of parameters. This simply does not have the same level of flexibility as having unrestricted control of the local volatility function.

Note that, following Derman and Kani (1998), it is possible to draw a link between stochastic volatility and local volatility. They show that the squared local volatility function that reproduces the same call prices as a given stochastic volatility model is

$$\sigma^2(K,T) = \mathbb{E}\left(\eta_T^2 \mid S_T = K\right).$$

In other words, if we replace the true underlying stochastic variance process with its conditional average, then call option prices remain unchanged.
1.5 Mixture Models

When constructing a local volatility model, the direct application of equation (1.3.1) requires knowledge of call option prices for all strikes and expiries. However, only a finite number of prices are observable in the market, meaning that a method for fitting a sufficiently differentiable curve to these prices is required to determine the local volatility function in practice. One such method, suggested by Brigo and Mercurio (2000), is to assume that the density of the stock price in the risk neutral measure is equal to the weighted average of a set of component densities. Each of these densities are generated by a simple component model, under which call options have an analytical price (e.g. the Black-Scholes model). Specifically, they let

$$\phi(x, t) = \sum_{k=1}^{n} \lambda_k \phi_k(x, t),$$

(1.5.1)

where $\phi(x, t)$ is the risk-neutral density of $S_t$, $\lambda_k$ is the mixture weight associated to component $k$, and $\phi_k(x, t)$ is the density of $S_{k,t}$ under the component model

$$dS_{k,t} = r_t S_{k,t} dt + \sigma_k(S_{k,t}, t) S_{k,t} dW_t.$$ 

Given a formula for the call price, $C_k(K, T)$, in component model $k$, the call pricing formula necessary to calibrate the mixture model is

$$C(K, T) = \int_{-\infty}^{\infty} (x - K)^+ \phi(x, t) dx$$

$$= \sum_{k=1}^{n} \lambda_k \int_{-\infty}^{\infty} (x - K)^+ \phi_k(x, t) dx$$

$$= \sum_{k=1}^{n} \lambda_k C_k(K, T)$$

Brigo and Mercurio’s main result is that the unique local variance function consistent with this mixture model is

$$\sigma^2(x, t) = \frac{\sum_{k=1}^{n} \lambda_k \sigma_k^2(x, t) \phi_k(x, t)}{\sum_{k=1}^{n} \lambda_k \phi_k(x, t)}$$

Thus, instead of using equation (1.3.1), they determine the local variance function by first calibrating a mixture model, and then taking the weighted average of the component’s local variance functions. The main advantages of this approach is that it avoids the need to specify an arbitrage-free interpolation between call prices, and guarantees that the resulting risk-neutral density of the stock price is well behaved.
However, Brigo and Mercurio also note that a mixture of standard Black-Scholes models with identical drifts is not sufficient to produce a skew in the implied volatility smile, in the sense that the minimum of the smile will always occur at the at-the-money strike. Therefore Brigo, Mercurio, and Sartorelli (2003) extend this approach to allow for component with differing drifts, i.e.

$$dS_{k,t} = \mu_{k,t}S_{k,t}dt + \sigma_k (S_{k,t}, t) S_{k,t}dW_t,$$

where $\mu_{k,t}$ is a time dependent drift parameter subject to the condition

$$S_0 = e^{-\int_0^T r_u du} E(S_T) = e^{-\int_0^T r_u du} \sum_{k=1}^n \lambda_k E(S_{k,T}) = \sum_{k=1}^n \lambda_k e^{\int_0^T (\mu_{k,u} - r_u) du} S_0.$$

The resulting local variance function is

$$\sigma^2(x, t) = \frac{\sum_{k=1}^n \lambda_k \sigma_k^2(x, t) \phi_k(x, t)}{\sum_{k=1}^n \lambda_k \phi_k(x, t)} + \frac{2 \sum_{k=1}^n \lambda_k (\mu_{k,t} - r_t) \int_x^\infty x \phi_k(x, t) dx}{x^2 \sum_{k=1}^n \lambda_k \phi_k(x, t)}.$$

As we will see via the numerical examples of chapters 4 and 6, a mixture of log-normal models with different drifts is very effective at fitting skewed implied volatility surfaces. Furthermore, the ability to choose time dependent (e.g. piecewise constant) drift and volatility parameters for each component means that the model can be calibrated one expiry at a time, which greatly reduces the computational burden.

An alternative way to fit skewed implied volatility surfaces is to start with component models that already allow for skew on their own, such as the shifted log-normal model suggested by Brigo and Mercurio (2002), or the hyperbolic-sine model suggested by Brigo et al. (2003). The drawback of this approach is that these models typically have fixed (i.e. non-time-dependent) parameters that determine the skew at all expiries. Therefore the mixture model has to be simultaneously calibrated to the entire implied volatility surface, instead of being calibrated one expiry at a time.

Due to its greater flexibility and ease of calibration, the mixture of log-normal models with different drifts may produce superior results to a mixture of skewed models with identical drifts. However, the extensions to stochastic interest rates and stochastic volatility we develop in chapters 4 and 5 are not compatible with differing drifts. Fortunately, in the case of stochastic volatility, the key parameters determining the skew (e.g. the volatility of volatility, and the instantaneous correlation between volatility and the stock price) are allowed to be time dependent functions, which means that the mixture model can still be calibrated one expiry at a time.
1.6 Stochastic Interest Rates

The main focus of this monograph is to extend current techniques in equity derivatives pricing to the case of stochastic interest rates. Of the common approaches to modelling interest rates, the two most popular are short-rate models and market models. The former model a single quantity, the current instantaneous interest rate, $r_t$, from which all other values need to be derived. The latter model an entire set of market observable quantities, namely forward rates or swap rates. The advantage of market models is that they can be easily and accurately calibrated to highly liquid interest rate derivatives, such as caps in the case of forward rate models, and swaptions in the case of swap rate models.

However, the high dimensionality and mathematical complexity of market models makes their incorporation into equity derivatives pricing quite difficult (see for example Grzelak, Borovykh, van Weeren, and Oosterlee (2008) and Grzelak and Oosterlee (2010), both of which make use of a number of approximations). Therefore, we choose to focus on the application of short-rate models. Following the presentation in Brigo and Mercurio (2007), a one-factor short-rate model has the form

$$dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dW_t,$$

where $W_t$ is a standard Brownian motion in the risk-neutral measure adapted to the filtration $\mathcal{F}_t$, and $\mu$ and $\sigma$ are sufficiently well-behaved functions to guarantee a unique strong solution for this SDE. The time $t$ price of a contract with pay-off $V_T$ at time $T > t$ is then

$$V_t = E\left(e^{-\int_t^T r_u du} V_T \middle| \mathcal{F}_t\right).$$

(1.6.1)

For example, the time $t$ price of a Zero Coupon Bond (ZCB) with expiry $T$, which has pay-off $V_T = 1$, is

$$P(t, T) = E\left(e^{-\int_t^T r_u du} \middle| \mathcal{F}_t\right).$$

In order to calibrate the model to the market yield curve, which is defined in terms of ZCB prices, we would like this expectation to have an analytical formula. After ZCBs, the next most important set of calibration instruments consists of caps and floors. These are essentially linear combinations of puts or calls on ZCBs, as explained in section 2.6.1 of Brigo and Mercurio (2007). Therefore, we would also like to have an analytical formula for the right hand side of equation (1.6.1) in the case that $V_T = (P(t, T) - K)^+$ and $V_T = (K - P(t, T))^+$.

Besides having efficient formulas for computing the prices of bonds, and options on bonds, there are a number of other important traits for a good short-rate model. For example, we may require that the short-rate is positive, mean-reverting, has finite variance, and can be easily simulated. This first requirement is perhaps not so important given recent experience of negative inter-
est rates in Europe and Japan. A popular example which satisfies all of these requirements is the Cox, Ingersoll Jr, and Ross (1985) model, which sets

$$\mu(r_t, t) = k(\theta - r_t), \quad \sigma(r_t, t) = \psi \sqrt{r_t},$$

with constant parameters $k, \theta, \psi \in \mathbb{R}^+$. Under this model, the short-rate has a non-central chi-squared distribution, and ZCB option prices have closed form expressions in terms of this distribution’s CDF. When choosing a short-rate model to combine with an equity model it is preferable that the resulting hybrid model is affine, so as to maintain analytical tractability. However, the CIR model typically leads to non-affine hybrid models because the instantaneous covariance between the log stock and short-rate will depend on the square-root of $r_t$.

The Hull and White (1990) model, on the other hand, leads to highly tractable hybrid models, and is therefore the short-rate model we focus on in this monograph. Nevertheless, many of our results may be applied to other short-rate models, even those with multiple factors. The drift and diffusion coefficients of the short-rate under the Hull-White model are

$$\mu(r_t, t) = k(\theta_t - r_t), \quad \sigma(r_t, t) = \psi,$$

with constant parameters $k, \psi \in \mathbb{R}^+$ and time-dependent parameter $\theta_t$. Note that, due to the time dependence in $\theta_t$, the model can be exactly calibrated to the market yield curve. Furthermore, a straight-forward extension to time dependent volatility, $\psi_t$, also allows the model to be exactly calibrated to at-the-money caps. Under this model the short-rate has a normal distribution, and ZCBs have a log-normal distribution, so that ZCB option prices are given by a Black-Scholes like formula. Moreover, the short-rate, bank-account, and ZCB prices can all be exactly simulated over arbitrary time steps, as they follow a simple transformation of the joint normal distribution. This makes the model highly suitable for Monte-Carlo simulation.

The advantage of the Hull-White model when constructing hybrid models can be seen through the example of the Black-Scholes Hull-White model, as presented by Brigo and Mercurio (2007). We have the bivariate SDE

$$dS_t = r_t S_t dt + \eta S_t dW_{1,t},$$
$$dr_t = k(\theta_t - r_t) dt + \psi dW_{2,t},$$

where $W_{1,t}$ and $W_{2,t}$ have correlation $\rho \in (-1, 1)$. Under this model, $\log(S_t)$, $r_t$ and $\log(B_t)$ are jointly normal, with explicitly known parameters, in both the risk-neutral and $T$-forward measures. This yields analytical formulas for vanilla option prices, and makes the model very easy to simulate. If we instead replace the short-rate with the CIR model, then closed form formulas are only available in the case that $\rho = 0$, and even then they take the form of integral expressions involving the characteristic function of the log stock.
1.7 Non-Parametric Pricing of Volatility Derivatives

The primary goal of the modelling approaches discussed above is to fit a set of parameters to the market prices of highly liquid derivatives. The prices of path dependent derivatives can then be determined using either analytical results if they are available, or Monte-Carlo simulation if not. However, in certain cases it is possible to determine a direct relationship between vanilla option prices and more exotic contracts.

One of the most important examples is that of volatility derivatives. These are contracts whose payoff depends on the observed quadratic variation, $\langle X \rangle_T$, of the log stock price, $X_t := \log (S_t)$. For instance, given a continuous stock price process of the form

$$dS_t = rS_t dt + \eta_t S_t dW_t,$$

where $\eta_t$ is some (possibly stochastic) volatility process, then the quadratic variation is

$$\langle X \rangle_T = \int_0^T \eta_t^2 dt.$$

We call $\langle X \rangle_T$ the “realized variance”, and its square-root, $\sqrt{\langle X \rangle_T}$, the “realized volatility”. Popular volatility derivatives include the variance swap, whose payoff is the realized variance minus a fixed strike, and the volatility swap, whose payoff is the realized volatility minus a fixed strike. Calls and puts on variance and volatility are also common.

In the case of a variance swap, Neuberger (1994) explains how to replicate the pay-off using a log contract, which pays $\log (S_T)$ at time $T$, and continuous trading of stocks. In particular, assuming that interest rates are deterministic and that the stock price process is continuous, he shows that the fair strike of a variance swap is

$$E (\langle X \rangle_T) = -2 \left( E (\log (S_T)) - \log \left( \frac{S_0}{P(0, T)} \right) \right),$$

where the expectation is taken in the risk neutral measure. Neuberger recommends that the log contract be traded so that it can be used to construct variance swaps and thus help hedge volatility. However, using the results of Breeden and Litzenberg (1978), it is possible to replicate a log contract using a static position in call and puts across the continuum of strikes. Of course, only a finite set of strikes are available in the market, but it is still possible to construct an approximation to the log contract in this case.

Building on this idea, Carr and Madan (1998) show how to replicate a number of volatility contracts by delta hedging a static option position. Importantly, their technique only requires the continuous trading of the underlying, not the continuous trading of options. In addition to the standard variance swap, they are able to replicate contracts paying the variance between two fu-
tutions, and contracts paying the variance over the period for which the futures price lies in a specified range. In a similar fashion to Neuberger (1994), they rely on the stock price process being continuous, and do not allow for stochastic interest rates.

More recently, Carr and Lee (2008) show how a much wider range of volatility derivatives, including calls and puts on variance and volatility, can be replicated using vanilla options. They first replicate the exponential variance contract, which pays $e^{\lambda \langle X \rangle_T}$, using the power contract, which pays $S_p^p$, for some constants $\lambda, p \in C$. They then explain how to construct a wide range of volatility options out of exponential variance contracts. As power contracts, just like log contracts, can be replicated using calls and puts, Carr and Lee are able to replicate general volatility derivatives via continuous trading in vanilla options, without having to fit a specific parametric model.

However, Carr and Lee make a number of strong assumptions, including that the underlying volatility process, $\eta_t$, and driving Brownian motion, $W_t$, are independent. Although this assumption enables them to replicate a much larger class of volatility derivatives than was previously possible, it is nevertheless problematic. Fortunately, they are able to extend their results so that they hold approximately even in the case of non-zero correlation between $\eta_t$ and $W_t$.

1.8 Outline of the Monograph

We begin in chapter 2 by reviewing existing approaches to pricing equity derivatives under stochastic interest rates, including local volatility, stochastic volatility, and the model-free pricing of variance swaps. Next, in chapter 3, we will see how to construct a flexible class of analytically tractable local volatility models under stochastic rates. All the models in this class have closed-form expressions for the joint density of the stock price, short-rate and bank account in both the risk-neutral and $T$-forward measures. They also allow vanilla options to be priced using a one-dimensional integral involving the normal CDF. We then develop a methodology for combining these models into mixture models in chapter 4. This allows us to accurately match the market implied volatility smile across multiple expiries.

In chapter 5 we extend this mixture model approach to include stochastic volatility, as this has a number of advantages over the pure local volatility model, including that it allows for more realistic evolution of the implied volatility curve, and more accurate prices for certain exotic derivatives. In contrast to the model based approaches of the previous chapters, we examine the non-parametric pricing of volatility derivatives in chapter 6. The results of that chapter are non-parametric in the sense that, conditional on a model for interest rates, we are able to relate the prices of a range of volatility derivatives directly to the market prices of call options, without assuming a specific parametric model for the stock price process. Finally, we summarise our contributions and conclude in chapter 7.
Chapter 2

Review of Equity Derivatives Pricing under Stochastic Interest Rates

In this chapter we will review existing work regarding the extension of local volatility, stochastic volatility, and non-parametric pricing to incorporate stochastic interest rates. In the case of model based methodologies (i.e. local volatility and stochastic volatility) we consider the following four criteria. Firstly, it should be possible to rapidly calibrate the model to the market prices of highly liquid instruments. This means that any necessary numerical algorithms should not be too computationally costly. Secondly, the model should be able to simultaneously reproduce the market prices of all of these calibration instruments. Thus it should closely fit the implied volatility surface across all strikes and expiries. Thirdly, the dynamics of the stock price, short-rate, and any other quantities of interest, should be as realistic as possible, and be consistent with empirical observations. Fourthly, the model should be easy to simulate, so that Monte-Carlo pricing is efficient.

Typically, the output of the calibration routine are the parameters governing the drift and diffusion coefficients of the underlying SDE. Once these coefficients are known, the SDE can be simulated using a small-time-step discretization scheme, such as the Euler scheme. This is fine when pricing derivatives whose payoff is dependent on the entire path, such as a barrier option. However, in the case that only a few points of the path need to be simulated, it is preferable that a simulation scheme which is accurate over long time steps is available.

On the other hand, non-parametric pricing methodologies are examined using a different set of criteria. Firstly, the underlying assumptions should not be too restrictive. Secondly, the methodology should be flexible enough to price a wide range of instruments, especially those of particular interest to practitioners. This second criteria is not a significant issue for model-based approaches
because, even if analytical results are not available, exotic derivatives can be priced using Monte-Carlo simulation. Note that, unlike model based approaches, non-parametric approaches write derivative prices directly in terms of the market prices of vanilla options, and are therefore automatically consistent with these prices. Conversely, model based approaches are only consistent with market prices if they are able to produce a good fit during calibration.

2.1 Local Volatility under Stochastic Interest Rates

One of the first papers to focus on local volatility under stochastic interest rates is Benhamou, Rivoira, and Gruz (2008). They extend Dupire’s formula for the local volatility function to allow for both stochastic interest rates and independent jumps in the stock price process. In the case of a continuous price process, which is our main concern, their formula reduces to

$$\sigma^2(K,t) = \frac{\partial}{\partial t} C - KP_0, t E_t \left( r_t I\{S_t > K\} \right) + \frac{y_t}{2} C - K \frac{\partial}{\partial K} C \frac{1}{2} K^2 \frac{\partial^2}{\partial K^2} C.$$  \hspace{1cm} (2.1.1)

where $E_t(\cdot)$ denotes the expectation in the $t$-forward measure. However, unlike Dupire’s formula, the right hand side cannot be computed directly from the market prices of liquid instruments. Specifically, the term $E_t \left( r_t I\{S_t > K\} \right)$ needs to be estimated conditional on a joint model for the stock price and short-rate.

In the case of Hull-White interest rates, Benhamou et al. (2008) examine the difference between the local volatility functions implied by a fixed surface of option prices before and after accounting for stochastic interest rates. They then develop an iterative algorithm for estimating this difference, and are thus able to calibrate their Local Volatility Stochastic Rates (LVSR) model. However, this algorithm is based on an approximation for the covariance between the log stock price and the short-rate. It is unclear how accurate this approximation is, and what effect different levels of instantaneous correlation, or increasing time to expiry, have on the results.

Grzelak et al. (2008) extend this approach to allow for a multi-factor short-rate process consistent with the stochastic volatility Libor market model. This has the advantage that the model is consistent with the smile in cap and swaption implied volatilities. However, the path of the short-rate may have very large discontinuities at the exercise dates associated to the forward rates. Their approach also has the same limitations as that of Benhamou et al. (2008) regarding the accuracy of the approximation for the correlation between the log stock price and the short-rate, as it relies on the same iterative algorithm to estimate the local volatility function.

Instead of computing formula (2.1.1) directly, an alternative approach is to develop formulas for the prices of vanilla options given a particular short-rate model and a parametric specification of the local volatility function. The necessary parameters can then be calibrated in the usual way, i.e. by minimizing the sum of squared differences between the model and market prices. For example, Benhamou, Gobet, and Miri (2012) derive an expansion formula for option prices with re-
2.1. Local Volatility under Stochastic Interest Rates

spect to a proxy model, specifically the time-dependent Black-Scholes model coupled with Gaussian interest rates. Unlike in the traditional local volatility model, they assume that volatility is a function of the stock price divided by the bank account. They also conduct some numerical experiments to demonstrate the accuracy of their approach under Hull-White interest rates and a (time-homogeneous) local volatility function of the form.

$$\sigma(x, t) = \nu x^{(\beta - 1)},$$

where $$\nu > 0$$ and $$\beta \in \mathbb{R}$$. Nevertheless, in the case of a time-inhomogeneous volatility specification, which is necessary to fit the implied volatility surface at multiple expiries, their formulas involve a number of high-dimensional integrals, and may be difficult to implement for more complex volatility specifications. Furthermore, the accuracy of the approximation may depend heavily on the similarity between the chosen model and the proxy model.

A third way to calibrate the local volatility function under stochastic interest rates is to use Monte-Carlo simulation. For instance, van der Stoep, Grzelak, and Oosterlee (2016) develop a method for applying equation (2.1.1) by estimating $$\mathbb{E}^t(r_t I\{S_t > K\})$$ efficiently. Under Hull-White interest rates, they begin by projecting the stock price, $$S_t$$, onto a standard normal random variable, $$X$$, using a technique known as stochastic collocation, which results in a function $$g(x)$$ such that $$S_t \overset{d}{=} g(X)$$. They then write the key expectation as an affine combination of the truncated moments of $$X$$, specifically

$$\mathbb{E}^t(r_t I\{S_t > K\}) = \left( \mu^t_r(t) + \sigma^t_r(t) \left( \sum_{k=0}^{n-1} \hat{\beta}_k \mathbb{E}^t(X^k | X > g^{-1}(K)) \right) \right) Q^t(S_t > K),$$

where $$\mu^t_r(t)$$ and $$\sigma^t_r(t)$$ are the mean and standard deviation of the short-rate in the $$t$$-forward measure. The necessary coefficients, $$\hat{\beta}_k$$, are estimated during the simulation using ordinary least squares (OLS) regression. Given a complete surface of call option prices, the calibration procedure proceeds as follows, starting from time 0:

1. Simulate forward one time step using a suitable discretization scheme, e.g. the Euler method.
2. Estimate $$\mathbb{E}^t(r_t I\{S_t > K\})$$ for the current time as a function of $$K$$ using the stochastic collocation and OLS regression based method.
3. Compute the local volatility function using equation (2.1.1).
4. Repeat steps 1, 2 and 3 until the final expiry is reached.

The accuracy of the calibration can be checked by comparing the model call prices, estimated using the simulated values of $$S_t$$, to the market call prices. Note that there are three potential sources of error. The first is due to the use of Monte-Carlo simulation. The second is due to the projection of the stock price onto a standard normal random variable. The third is due to the projection of the key conditional expectation onto the truncated moments of this standard normal random variable.
Instead of using Monte-Carlo simulation, it is also possible to calibrate the local volatility function using a PDE based method. As explained by Ren, Madan, and Qian (2007), this involves solving the Fokker-Plank equation forward in time to determine the joint density of the stock price and short rate. At each time step, $E^t \left( r_t \mathbb{1}_{S_t > K} \right)$ can be computed from this joint density and then fed into equation (2.1.1) in order to compute the local volatility function needed to move on to the next time step. Nevertheless, this procedure requires the numerical solution of a two-dimensional second-order PDE, which is quite computationally costly.

Overall, the existing methods for calibrating a local volatility function under stochastic interest rates are limited by either significant computational burden or the need for multiple approximations. This is what motivates our development in chapters 3 and 4 of two highly tractable methods for calibrating the local volatility function which rely on no approximations, and provide simple formulas of vanilla option prices in terms of one-dimensional integrals.

## 2.2 Stochastic Volatility under Stochastic Interest Rates

As explained in section 1.3, stochastic volatility is often preferred to local volatility because it implies more realistic dynamics for stock prices and the implied volatility surface. The most popular approach to constructing stochastic volatility models under stochastic interest rates is the same as that used under deterministic interest rates. The key idea is to make the drift coefficient and instantaneous covariance matrix an affine function of the state variables, which are in this case the stock-price, short-rate and volatility. As explained in section 1.4, this means that the characteristic function of the log stock price can be found using the results of Duffie et al. (2000), and vanilla options can be priced using the FFT based techniques of Carr and Madan (1999).

A straightforward example is the Schöbel-Zhu-Hull-White (SZHW) model presented by van Haastrecht, Lord, Pelsser, and Schrager (2009). Under this model the short-rate and volatility each follow an Ornstein-Uhlenbeck process, i.e.

\[
\begin{align*}
    dS_t &= r_t S_t dt + \eta_t S_t dW_{1,t}, \\
    dr_t &= (\theta_t - \alpha r_t) dt + \psi dW_{2,t}, \\
    d\eta_t &= \kappa (\bar{\eta} - \eta_t) dt + \gamma dW_{3,t},
\end{align*}
\]

where $\gamma$ is the volatility of volatility, $\bar{\eta}$ is the long-run average volatility, $\kappa$ is the mean reversion rate, and $(W_1, W_2, W_3)$ is a correlated joint Brownian motion in the risk-neutral measure. Importantly, this model can be made affine by replacing $S_t$ with $X_t := \log (S_t)$ and adding $\eta^2_t$ as a fourth state variable. Furthermore, the coupled ODEs needed to derive the characteristic function all have analytical solutions, except for one, whose solution is nevertheless available in terms of the ordinary hypergeometric function. Thus, the characteristic function can be rapidly evaluated, which makes calibration not too difficult.

Similarly, Grzelak and Oosterlee (2011) extend the Heston stochastic volatility model with each
of Hull-White and CIR interest rates. However, neither of the resulting models is affine when all
the driving Brownian motions have non-zero correlations with each other. Therefore, they instead
impose a correlation between the short rate and the other two processes by using a model of the form
\[
\begin{align*}
    dS_t &= r_t S_t dt + \sqrt{v_t} S_t dW_{1,t} + \Omega_t r^p_t S_t dW_{2,t} + \Delta \sqrt{v_t} S_t dW_{3,t}, \\
    dr_t &= (\theta_t - ar_t) \, dt + \psi r^p_t dW_{2,t}, \\
    dv_t &= \kappa (\bar{v} - v_t) \, dt + \gamma \sqrt{v_t} dW_{3,t},
\end{align*}
\]
where \( p = 0 \) for the Hull-White model, \( p = \frac{1}{2} \) for the CIR model, and all the Brownian motions are
independent except for the pair \((W_1, W_2)\). Here the coefficients \( \Omega_t \) and \( \Delta \) give indirect control over
the correlation between the state variables. In contrast to the SZHW model, in which the correlation
matrix can be set exactly, Grzelak and Oosterlee (2011) rely on an approximation for \( \Omega_t \) in order to
achieve a target level of correlation.

A second approach to calibrating the Heston Hull-White (HHW) model, presented by Antonov,
Arneguy, and Audet (2008), is to use the technique of Markovian projection to approximate the non-
affine model with a more tractable process. In particular, they project the model onto a shifted
Heston model with displaced volatility, which is affine and can thus be handled using standard
techniques. Nevertheless, as noted by van Haastrecht (2010), this approximation may break down
for extreme parameter values and long expiries.

## 2.3 Combined Local and Stochastic Volatility

For all the models presented in the previous section, the majority of the parameters are assumed
to be constant so as to simplify the derivation of the characteristic function. Although extensions
to time dependent parameters are possible, this still does not yield the same level of flexibility as
the local volatility model when it comes the reproducing the market implied volatility surface. For
this reason it is quite common to combine both Local and Stochastic Volatility into a single model
(LSV).

However most work to date assumes deterministic interest rates. For example: An and Li (2015)
and Lorig, Pagliarani, and Pascucci (2015) develop asymptotic expansions for option prices; Lipton,
and van der Stoep et al. (2016) utilize Monte-Carlo simulation; and Piterbarg (2007) employs an
approximation using Markovian projection. Note that in some ways the problem of calibrating an
LSV model under deterministic interest rates is similar to that of calibrating a local volatility model
under stochastic interest rates. To see this, consider the model proposed by Ren et al. (2007),
\[
\begin{align*}
    dS_t &= r_t S_t dt + \sigma (S_t, t) \eta_t dW_{1,t} \\
    d\eta_t &= \mu_\eta (\eta_t, t) \, dt + \sigma_\eta (\eta_t, t) \, dW_{2,t}.
\end{align*}
\]
Also let $\sigma_{LV}(K,t)$ be the local volatility function given by Dupire’s formula, i.e. without a stochastic volatility factor. Then it can be shown that

$$\sigma^2(K,t) = \frac{\sigma^2_{LV}(K,t)}{\mathbb{E}^t(\eta_t \mid S_t = K)}.$$ 

This model can be calibrated by solving the Fokker-Plank equation forwards in time to get the joint density of $S_t$ and $\eta_t$. At each time step the term $\mathbb{E}^t(\eta_t \mid S_t = K)$ is computed using numerical integration and entered into the above formula, which yields the current local volatility factor necessary to move on to the next step. This process is very similar to the PDE based calibration of a Local Volatility Stochastic Rates (LVSR) model, except that the second state variable is $\eta_t$, instead of $r_t$, and at each step we need to compute $\mathbb{E}^t(\eta_t \mid S_t = K)$ instead of $\mathbb{E}^t(r_t \mathbb{1}_{S_t > K})$. Likewise, the Monte-Carlo method proposed by van der Stoep et al. (2016) for calibrating LVSR models using stochastic collocation works in more or less the same way when calibrating LSV models.

All of the above mentioned methods share similar limitations to those proposed for calibrating LVSR models, namely computational difficulty or the need for approximations that may cause problems in certain cases. Alternatively, Ramponi (2011) extend the mixture model approach of Brigo and Mercurio (2000) to allow for stochastic volatility and regime switching models. As an example, they propose a mixture of Heston models, which combines the realistic volatility dynamics provided by the square-root process, with the ability of a local volatility model to match the implied volatility surface. Nevertheless, their model does not include stochastic interest rates, and they do not provide an efficient method for computing the joint density of the log stock price and volatility under each component model, which is necessary in order to simulate the process.

One of the few papers to incorporate stochastic interest rates into a LSV framework is Deelstra and Rayée (2013). Although they consider an FX context, in which both the foreign and domestic interest rates are stochastic, their model can be written as follows for an equity context:

$$dS_t = r_t S_t dt + \sigma(S_t, t) \phi(\nu_t, t) S_t dW_{1,t}$$
$$dr_t = (\theta_t - a_t r_t) dt + \psi_t dW_{2,t},$$
$$d\nu_t = \alpha(\nu_t, t) dt + \beta(\nu_t, t) dW_{3,t}$$

where $\sigma(S_t, t)$ is the local volatility factor, and $\phi(\nu_t, t)$ is the stochastic volatility factor, which depends on some stochastic process, $\nu_t$. Deelstra and Rayée begin by assuming that the pure LVSR model has been calibrated, i.e. the function $\tilde{\sigma}(x, t)$ in the model

$$dS_t = r_t S_t dt + \tilde{\sigma}(S_t, t) S_t dW_{1,t}$$
$$dr_t = (\theta_t - a_t r_t) dt + \psi_t dW_{2,t},$$

is taken as given. They then show that the local volatility factor, that reproduces the same option
prices as this LVSR model, satisfies
\[
\sigma^2(K, t) = \frac{\tilde{\sigma}^2(K, t)}{\mathbb{E}^t(\phi^2(\nu_t, t)|S_t = K)}.
\]

(2.3.1)

The function \( \sigma(K, t) \) can then be determined using a PDE based approach, which involves solving for the joint density of \((S_t, r_t, \nu_t)\) in the \( t \)-forward measure. However, numerically solving a three-dimensional second-order PDE may be quite computationally difficult, and Deelstra and Rayée (2013) provide no numerical examples or calibration experiments. Note that, in the case of independent volatility, the term \( \mathbb{E}^t(\phi^2(\nu_t, t)|S_t = K) = \mathbb{E}^t(\phi^2(\nu_t, t)) \) can be computed analytically, thus avoiding the need to solve this PDE numerically. However, as \( \nu_t \) appears in the SDE governing \( S_t \), the assumption that \( \nu_t \) and \( S_t \) are independent is unrealistic, even if they are instantaneously uncorrelated.

Note that, in chapters 3 and 4, we are concerned with calibrating the local volatility function under stochastic rates, i.e. finding \( \tilde{\sigma}^2(K, t) \), meaning that equation (2.3.1) is not relevant there. In chapter 5, we go on to calibrate the local volatility function under stochastic volatility and stochastic rates. However, the condition that \( \nu_t \) and \( S_t \) are independent is not satisfied, so equation (2.3.1) does not help us simplify the problem.

### 2.4 Volatility Derivatives under Stochastic Interest Rates

Although much work has been done regarding the non-parametric pricing of volatility derivatives, very little is known in the context of stochastic interest rates. One of the few papers to deal with this issue is Hörfelt and Torné (2010). They consider a set-up where the stock price and zero-coupon bonds follow continuous processes with some unknown stochastic volatilities. Fixing the final expiry \( T \), and writing \( P_t := P(t, T) \) for the unit zero-coupon bond, their model is

\[
dS_t = r_t S_t dt + \sigma_{S,t} S_t dW_{1,t},
\]
\[
dP_t = r_t P_t dt + \sigma_{P,t} P_t dW_{2,t},
\]

where \( r_t, \sigma_{S,t} \) and \( \sigma_{P,t} \) are themselves stochastic process. Switching to the \( T \)-forward measure, Hörfelt and Torné show that

\[
\int_0^T \frac{dS_t}{S_t} = -\log(P_0) + \int_0^T \left( \rho \sigma_{P,t} \sigma_{S,t} - \frac{1}{2} \sigma_{P,t}^2 \right) dt + \int_0^T \sigma_{S,t} dW_{1,t}^T - \int_0^T \sigma_{P,t} dW_{2,t}^T,
\]

where \( W_{1,t}^T \) and \( W_{2,t}^T \) are Brownian motions with correlation \( \rho \). Furthermore, applying Ito's formula to \( \log(S_t) \) and integrating the result yields the well known formula

\[
\int_0^T \frac{dS_t}{S_t} = \log\left( \frac{S_T}{S_0} \right) + \frac{1}{2} \int_0^T \sigma_{S,t}^2 dt.
\]
Combining these equations and taking the expectation of both sides, Hörfelt and Torné find that the annualized fair strike of a variance swap is

\[
\nu_{vs} = \mathbb{E}^T \left( \frac{1}{T} \int_0^T \sigma_{S,t}^2 dt \right) \\
= -\frac{2}{T} \mathbb{E}^T \left( \log \left( \frac{S_T}{S_0} \right) \right) - \frac{2}{T} \log (P_0) + \mathbb{E}^T \left( \frac{1}{T} \int_0^T \left( 2\rho \sigma_{P,t} \sigma_{S,t} \sigma_{P,t} - \sigma_{S,t}^2 \right) dt \right).
\]

Importantly, the first two terms on the final line are fully determined by the prices of vanilla options and the current yield curve. Thus, defining

\[
\nu_{vs}^* = -\frac{2}{T} \mathbb{E}^T \left( \log \left( \frac{S_T}{S_0} \right) \right) - \frac{2}{T} \log (P_0)
\]

to be the fair strike of a variance swap under deterministic interest rates, and

\[
\nu_{bus} = \mathbb{E}^T \left( \frac{1}{T} \int_0^T \sigma_{P,t}^2 dt \right)
\]

to be the fair strike of a variance swap on a bond, then the effect of stochastic interest rates is

\[
\nu_{vs} - \nu_{vs}^* = \mathbb{E}^T \left( \frac{2}{T} \int_0^T \rho \sigma_{P,t} \sigma_{S,t} dt \right) - \nu_{bus}.
\]

Thus, if \( \rho \leq 0 \) then the fair strike needs to be adjusted downwards compared to the deterministic interest rates case, whereas if \( \rho > 0 \) then the direction of the adjustment is uncertain. Furthermore, the adjustment will tend to be larger for longer-term volatility derivatives, as the volatility of bonds typically increases with time to expiry. Using the Cauchy-Schwartz inequality, Hörfelt and Torné are also able to place the following bounds on \( \sqrt{\nu_{vs}} \) for \( \rho \geq 0 \),

\[
\sqrt{\nu_{vs}^*} - \nu_{bus} \leq \sqrt{\nu_{vs}} \leq \sqrt{\nu_{vs}^*} - \left(1 - \rho^2\right) \nu_{bus} + \rho \sqrt{\nu_{bus}},
\]

with the inequalities reversed for \( \rho \leq 0 \). They also note that, if \( \sigma_{S,t} \) and \( \sigma_{P,t} \) are fixed constants, then the right inequality becomes an equality, and can be used to approximate \( \sqrt{\nu_{vs}} \) under stochastic interest rates. In fact, using an example equity-interest-rate hybrid model calibrated to S&P index option data, they find that this approximation performs quite well in realistic scenarios.

However, this analysis has a couple of limitations. Firstly, it only applies to variance swaps, and cannot be used to value other types of volatility derivative. Secondly, the term \( \mathbb{E}^T \left( \frac{2}{T} \int_0^T \rho \sigma_{P,t} \sigma_{S,t} dt \right) \) needs to be approximated if a specific model for stock and bond prices is not available. Thus, exact non-parametric results are limited to the case that \( \rho = 0 \).
2.5 Conclusion

Although the design and calibration of equity-interest-rate-hybrid models has received much attention in recent literature, it still poses a number of challenges. Existing methods often rely on the use of numerical PDEs, Monte-Carlo simulation, asymptotic expansions or Markovian projections. These methods are frequently numerical intensive, dependent on approximations that may breakdown in certain situations, or restricted to a limited set of models.

On the other hand, affine models are a tractable and exact way to combine stochastic volatility and stochastic rates, but lack the local volatility component necessary to reproduce the market implied volatility surface. Similarly, mixture models are an effective way to combine local and stochastic volatility, but do not extend to stochastic rates. Throughout chapters 3 and 4 we develop highly tractable and exact methods for calibrating local volatility models under stochastic interest rates, and then extend this to incorporate stochastic volatility in chapter 5.

Instead of relying on the above model based approaches, volatility derivatives can also be priced by relating them directly to the prices of vanilla options. Nevertheless, in the context of stochastic interest rates, results are only known for the basic variance swap. Thus, in chapter 6, we extend non-parametric pricing under stochastic interest rates to include a variety of other volatility derivatives.
Chapter 3

Parametric Local Volatility Models under Stochastic Interest Rates

3.1 Introduction

In the previous chapter we saw that existing methods for determining the local volatility function under stochastic interest rates are limited by their reliance on costly numerical methods, or potentially unreliable approximations. Therefore, in this chapter, we will introduce a new methodology for constructing analytically tractable LVSR models without either of these drawbacks. It is based on the observation by Carr, Tari, and Zariphopoulou (1999) that a valid arbitrage-free model can be created by equating the stock price to the conditional expectation of a function of the underlying Brownian motion. We extend this approach to stochastic interest rates, and present a new example that guarantees positive stock prices and is able to reproduce the extreme skew often observed in implied volatility smiles. The resulting LVSR model allows call options to be priced via a one-dimensional numerical integration.

Such models are useful in their own right, as they account for both implied volatility skews and stochastic interest rates. Furthermore, the joint distribution of the stock price, short-rate, and bank account for these models is known analytically in terms of the multivariate normal distribution, meaning that they can be simulated exactly over arbitrarily long time-steps. However, given only a single time-dependent volatility parameter, they may be unable to reproduce the entire implied volatility surface across all strikes and expiries. Fortunately, as we will see in the following chapter, this limitation can be overcome by using the models developed here as components in a mixture model.
Importantly, the ability of a mixture model to generate an implied volatility skew depends on that of the underlying component models. In particular, under a mixture of standard Black-Scholes models, the minimum of the implied volatility smile will always occur at the at-the-money (ATM) strike. However, real world option prices often display a downwards sloping skew, meaning that the implied volatility consistently decreases with strike, rather than reaching a minimum at the ATM level. Thus, component models based on the Black-Scholes models are insufficient to produce accurate mixture models. This makes the LVSR models developed in this chapter especially useful as components in the mixture model approach studied in the next chapter.

Our approach is in some ways similar to the hybrid Markov-functional model proposed by Fries and Eckstaedt (2011) in an FX context. Under stochastic domestic interest rates, they write the FX rate as an arbitrary time-dependent function of a driving diffusion process. However, this requires that the drift of the driving process is fully state dependent, meaning that the density of the FX rate is not known analytically. Furthermore, this drift needs to be determined for each point on a two dimensional grid by numerically solving a particular no arbitrage equation, which itself involves a two dimensional numerical integral. This problem is compounded by the fact that the drift needs to be recomputed during every iteration of the calibration routine before vanilla options can be priced. We are able to avoid these problems by modelling the stock price divided by the bank account instead of directly modelling the stock price, and by restricting the choice of the functional form so that it automatically satisfies the no arbitrage condition.

In section 3.2 below, we begin by describing a general framework for constructing LVSR models, and explaining how vanilla options can be efficiently priced under this framework. We then examine three specific examples and compare their theoretical characteristics in section 3.3. These examples are calibrated to real market data in section 3.4 in order to demonstrate their ability to reproduce the implied volatility smile. Finally, we conclude in section 3.5.

### 3.2 Assumptions and Main Results

The primary goal of this section is to write the stock price as a function of two jointly normal random variables, one depending on the stock’s driving Brownian motion, and the other depending on the short-rate’s driving Brownian motion. We can then price vanilla options by integrating their discounted payoffs against a bivariate normal density. Furthermore, in the case of Hull-White interest rates, we are able to reduce this computation to a one-dimensional integral involving the standard normal cumulative distribution function. We begin by detailing our basic framework in the assumption below.

**Assumption 3.2.1.** Assume that markets are frictionless and that there exists an equivalent risk-neutral measure, \( \mathbb{Q} \), such that the stock price, short-rate and bank account follow diffusive processes of the form

\[
    dS_t = (r_t - y_t)S_t dt + \eta(S_t, B_t, t) S_t dW_{1,t},
\]
3.2. Assumptions and Main Results

\[ dr_t = \mu_r(r_t, t)dt + \sigma_r(r_t, t)dW_{2,t}, \]
\[ dB_t = r_tB_t dt, \]

where \((W_{1,t}, W_{2,t})\) is a two-dimensional \(\mathcal{F}_t\)-adapted \(\mathbb{Q}\)-Brownian motion with time dependent correlation \(\rho_t \in (-1, 1)\), and \(y_t\) is a deterministic time-dependent dividend rate. The drift coefficient, \(\mu_r : \mathbb{R} \times [0, \infty) \to \mathbb{R}\), diffusion coefficient, \(\sigma_r : \mathbb{R} \times [0, \infty) \to \mathbb{R}\), and volatility function, \(\eta : \mathbb{R}^3 \times [0, \infty) \to \mathbb{R}\), must be measurable functions such that \((S_t, r_t, B_t)\) has a unique strong solution.

Our approach is a little non-standard in that we allow the local volatility function to depend on the bank account, as well as the stock price and time. Although we are interested in writing \(S_t\) as a function of normal random variables, the presence of the stochastic short-rate in its drift makes it challenging to work directly with \(S_t\). Therefore, we instead work with the “adjusted” stock price, \(X_t\), as defined below.

**Definition 3.2.1.** Let the adjusted stock price be \(X_t := S_tD_t^{-1}\), where the adjustment factor, \(D_t\), is

\[ D_t := A_tB_t, \quad A_t := e^{-\int_0^t y_udu}, \]

with bank account, \(B_t\), and dividend rate, \(y_t\), as in assumption 3.2.1.

Applying Itô’s rule to \(X_t\) yields

\[
dX_t = S_t (y_t - r_t) D_t^{-1} dt + D_t^{-1} ((r_t - y_t)S_t dt + \eta(S_t, B_t, t)S_t dW_{1,t})
\]
\[
= \eta(S_t, B_t, t)X_t dW_{1,t}
\]

(3.2.1)

We see that our adjustment to the stock price has removed the influence of the interest rate and the dividend rate on the drift. We now proceed by writing \(X_t\) as a function of the underlying Brownian motion and time as follows.

**Assumption 3.2.2.** Assume that there exists a twice-differentiable function, \(f : \mathbb{R} \times [0, \infty) \to \mathbb{R}\), such that the adjusted stock price, \(X_t\), satisfies

\[ X_t = f(Y_t, t), \quad Y_t := \int_0^t \nu_udW_{1,u}, \quad f(0, 0) = S_0, \]

where \(\nu_t \in \mathbb{R}\) is a deterministic function of time, and \(f(y, t)\) is strictly increasing in \(y\) for all \(t \geq 0\). We call \(f(Y_t, t)\) the “stock pricing function”.

As explained in the introduction, the above assumption is essentially a form of Markov Functional (MF) model. This type of model, which was first introduced in an interest rate context by Hunt, Kennedy, and Pelsser (2000), involves writing the economic variables of interest (e.g. bond prices, stock prices, or foreign currency prices) as time dependent functions of an underlying low-dimensional Markov process. In the case of equity modeling, Fries (2006) are able to calibrate a MF
model to a given surface of vanilla option prices. Although it is also possible to allow for stochasticity in the interest rate process by choosing the drift of the driving Markov process, the fact that this driver is one-dimensional makes it difficult to control the joint distribution of the short-rate and stock price under their approach.

Here we have introduced the parameter $\nu_t$ so that the driving Gaussian process, $Y_t$, may have time dependent volatility. Also, we assume that $f(y, t)$ is increasing in $y$ so that it has inverse function with respect to $y$, i.e. we can write $Y_t = f^{-1}(X_t, t)$. Applying Itô’s rule to $f(Y_t, t)$ yields

$$dX_t = \frac{\partial f}{\partial t}(Y_t, t)dt + \frac{\partial f}{\partial y}(Y_t, t)\nu_t dW_{1,t} + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(Y_t, t)\nu_t^2 dt.$$ 

Comparing this to equation (3.2.1) yields the fundamental PDE governing the function $f$,

$$\frac{\partial f}{\partial t}(y, t) + \frac{1}{2} \frac{\partial^2 f}{\partial y^2}(y, t)\nu_t^2 = 0. \tag{3.2.2}$$

Thus, if we are able to find a function that satisfies this PDE, we can use it to generate an arbitrage free LVSR model by setting $S_t = D_t f(Y_t, t)$. The resulting local volatility function is

$$\eta(S_t, B_t, t) = \nu_t X_t^{-1} \frac{\partial f}{\partial y}(f^{-1}(X_t, t), t). \tag{3.2.3}$$

Observe that the volatility function depends on $S_t$ and $B_t$ through $X_t$ only. Thus it will depend on $B_t$ if and only if it also depends on $S_t$. In other words, our framework is restricted to the set of models in which volatility depends on $S_t/B_t$ (and time) only. This contrasts with the volatility functions typically found in the literature, which depend on $S_t$ only. Also note that the reason we require $f(y, t)$ to have an inverse with respect to $y$ is so that the volatility term, $\eta$, can be written as a function of the state variables $S_t$ and $B_t$. If this was not the case, then $Y_t$ would have to be included as an additional state variable in the model.

Essentially, by choosing to model $X_t$ instead of $S_t$, and letting the volatility function depend on $X_t$, we have reduced the problem to the case of zero interest rates. Thus, our stock pricing function is governed by the same conditions as found in Carr et al. (1999), assuming that the constant interest rate and dividend rate in their model are set to zero. However, the pricing of vanilla options is slightly more complex because we are dealing with a two-dimensional process. In particular, the time zero price, $C(K, t)$, of a call option with strike $K$ and expiry $t$ satisfies

$$\frac{C(K, t)}{P_{0,t}} = \mathbb{E}^t \left( (S_t - K)^+ \right)$$

$$= \mathbb{E}^t \left( (D_t f(Y_t, t) - K)^+ \right), \tag{3.2.4}$$

where $\mathbb{E}^t(\cdot)$ denotes the expectation in the $t$-forward measure, and $P_{0,t}$ is the time 0 price of a zero-coupon bond paying $1$ at time $t$. In order to evaluate this expectation we first need to choose a model for the short-rate such that the joint distribution of $Y_t$ and $B_t$ can be determined analytically.
This requirement is the main factor that limits the class of interest rate models compatible with our approach. A convenient choice, which we will adopt throughout this chapter, is the Hull-White short-rate model, whose details are given in the assumption below.

**Assumption 3.2.3.** Under assumption 3.2.1, let the short-rate follow an Ornstein-Uhlenbeck process with drift and diffusion coefficients

\[ \mu_r(r_t, t) = \theta_t - a_t r_t, \quad \sigma_r(r_t, t) = \psi_t, \]

where \(\theta_t, a_t\) and \(\psi_t\) are deterministic functions of \(t\).

This model, also known as the extended Vasicek model, is well studied in the literature, and we refer the reader to Gurrieri, Nakabayashi, and Wong (2009) for details regarding the case of fully time-dependent parameters in the risk-neutral measure. The exact joint distribution of the short-rate and the bank account in this measure has also been given by Fries (2016), and by Ostrovski (2013), who are both concerned with deriving efficient Monte-Carlo schemes.

However, the joint distribution of the short-rate and the bank account in the \(t\)-forward measure does not appear to have been written down previously in the case of time dependent parameters, so we include it here for completeness. Note that the measure change is deterministic, so that the form of the joint distribution is the same as in the risk-neutral measure, just with different parameters. This is shown for the case of time-independent parameters in Brigo and Mercurio (2007) section 12.1.1.

**Theorem 3.2.1.** Under assumptions 3.2.1 and 3.2.3, define the functions

\[ h(u, v) = \psi_u e^{-\int_u^v a_y dy}, \quad H(u, t) = \int_u^t h(u, v) dv. \]

Then \(r_t\) and \(Z_t := \log (B_t)\) have a bivariate normal distribution in the \(t\)-forward measure, with parameters

\[ \mathbb{E}^t (r_t) = f_{mkt}(0, t), \quad \mathbb{E}^t (Z_{k,t}) = -\log (P_{0,t}) - \frac{1}{2} \int_0^t H^2(u, t) du, \]

\[ \operatorname{Var}^t (r_{k,t}) = \int_0^t h^2(u, t) du, \quad \operatorname{Var}^t (Z_{k,t}) = \int_0^t H^2(u, t) du, \]

\[ \operatorname{Cov}^t (r_{k,t}, Z_{k,t}) = \int_0^t h(u, t) H(u, t) du. \]

A detailed derivation of this theorem is given in appendix 3.A. Note that each integral above has an analytical solution for piecewise constant parameters. Now that we have established our
chosen interest rate model, we can detail the joint distribution of \( r_t, Z_t \) and \( Y_t \). See appendix 3.A for the proof.

**Theorem 3.2.2.** Under assumptions 3.2.1 to 3.2.3 the processes \( r_t, Z_t \) and \( Y_t \) have a joint normal distribution in the \( t \)-forward measure, with parameters

\[
\mathbb{E}^t(Y_t) = -\int_0^t \nu_u \rho_u H(u, t) du, \quad \text{Var}^t(Y_t) = \int_0^t \nu_u^2 du,
\]

\[
\text{Cov}^t(Y_t, r_t) = \int_0^t \nu_u \rho_u h(u, t) du, \quad \text{Cov}^t(Y_t, Z_t) = \int_0^t \nu_u \rho_u H(u, t) du,
\]

Again, these integrals have analytical solutions for piecewise constant parameters. We now have all the pieces necessary to price a call options using equation (3.2.4).

**Theorem 3.2.3.** Under assumptions 3.2.1 to 3.2.3, let \( \phi_Y(y) \) be the PDF of \( Y_t \), \( N(x) \) be the standard normal CDF, and

\[
a := \inf \{ y : f(y, t) > 0, y \in \mathbb{R} \}, \quad b(y) := \log \left( \frac{K}{A_t f(y, t)} \right), \]

\[
\hat{\mu}(y) := \mathbb{E}^t(Z_{k,t} | Y_{k,t} = y) = \mathbb{E}^t(Z_{k,t}) + \frac{\text{Cov}^t(Y_{k,t}, Z_{k,t})}{\text{Var}^t(Y_{k,t})} (y - \mathbb{E}^t(Y_{k,t})) , \]

\[
\hat{\sigma}^2 := \text{Var}^t(Z_{k,t} | Y_{k,t} = y) = \text{Var}^t(Z_{k,t}) - \frac{\text{Cov}^t(Y_{k,t}, Z_{k,t})^2}{\text{Var}^t(Y_{k,t})}.
\]

Then the price of a call option satisfies

\[
\frac{C(K, t)}{P_{0,t}} = \int_a^\infty \left( A_t f(y, t) e^{\hat{\mu}(y) + \frac{1}{2} \hat{\sigma}^2} N \left( \frac{\hat{\mu}(y) - b(y) + \hat{\sigma}^2}{\hat{\sigma}} \right) - K N \left( \frac{\hat{\mu}(y) - b(y)}{\hat{\sigma}} \right) \right) \phi_Y(y) dy, \quad (3.2.5)
\]

**Proof.** We write the call price as a two-dimensional integral, and then evaluate the inner integral by writing it in terms of the conditional distribution of \( Z_t \) given \( Y_t \).

\[
\frac{C(K, t)}{P_{0,t}} = \mathbb{E}^t \left( (A_t e^{Z_t} f(Y_t, t) - K)^+ \right)
\]

\[
= \int_a^\infty \int_b^\infty (A_t e^{Z} f(y, t) - K) \phi_{Y,Z}(y, z) dz dy
\]

\[
= \int_a^\infty \left( A_t f(y, t) \int_b^\infty e^{Z} \phi_{Y,Z}(y, z) dz - K \int_b^\infty \phi_{Y,Z}(y, z) dz \right) \phi_Y(y) dy
\]

\[
= \int_a^\infty \left( A_t f(y, t) e^{\hat{\mu}(y) + \frac{1}{2} \hat{\sigma}^2} N \left( \frac{\hat{\mu}(y) - b(y) + \hat{\sigma}^2}{\hat{\sigma}} \right) - K N \left( \frac{\hat{\mu}(y) - b(y)}{\hat{\sigma}} \right) \right) \phi_Y(y) dy,
\]

\[
(3.2.6)
\]
Note that $a$ is simply $f^{-1}(0, t)$, unless $f(y, t) > 0$ for all $y \in \mathbb{R}$, in which case $f^{-1}(0, t)$ is not defined and $a = -\infty$.

This theorem is the key result that allows us to rapidly price call options under the Hull-White model, and thus calibrate our LVSR model. In general, any short-rate model under which the joint distribution of $Y_t$ and $Z_t$ is available in the $t$-forward measure can be used to price call options using equation (3.2.6), but it may not be possible to reduce it to a one-dimensional integral as done above. Furthermore, if we must use a two-dimensional integral to evaluate option prices, it is possible to do so in the risk-neutral measure by using the equation

$$
C(K, t) = \mathbb{E} \left( B_t^{-1} (S_t - K)^+ \right)
= \mathbb{E} \left( e^{-Z_t} (A_t e^{Z_t} f(Y_t, t) - K)^+ \right).
$$

In this case, it is not necessary to change to the $t$-forward measure at all, and the joint distribution of $Y_t$ and $Z_t$ is only required in the risk-neutral measure.

When producing an example calibration for section 3.4, the integral in equation (3.2.5) is computed using Simpson’s rule with 100 intervals. Note that the payoff is positive if and only if

$$
Y_{k,t} > f^{-1} \left( KA_t^{-1} e^{-Z_{k,t}} \right).
$$

Therefore, using an integration range of 5 standard deviations above and below the mean, we set the upper and lower bounds to

$$
b_{\text{lower}} = \max \left( f^{-1} \left( KA_t^{-1} e^{-E'(Z_{k,t}) - 5\sigma'(Z_{k,t})} \right) , E'(Y_{k,t}) - 5\sigma'(Y_{k,t}) \right),
\quad b_{\text{upper}} = b_{\text{lower}} + 10\sigma'(Y_{k,t}),
$$

where $\sigma'(\cdot)$ denotes the standard deviation in the $t$-forward measure. This choice of integration range and number of intervals gave call prices that were accurate to 5 significant figures for all the strikes available in the market.

Note that equation (3.2.5) is also used to calibrate the mixture models presented in chapter 4, as it is need to price calls in each component model. However, in chapter 5, we use a one-dimensional fractional FFT to evaluate option prices under stochastic volatility and stochastic rates, meaning that equation (3.2.5) is not involved.

The next step in our program is find a solution to PDE (3.2.2). Following Carr et al. (1999), we first choose some final time horizon, $T$, and set the terminal condition

$$
f(y, T) = g(y) \forall y \in \mathbb{R}.
$$
Then the Feynman-Kac formula yields the following solution for $t \in [0, T]$:

$$f(y, t) = \mathbb{E}(g(Y_T)|Y_t = y). \tag{3.2.7}$$

Here $Y_T|Y_t = y$ is normally distributed with mean $y$, and variance

$$\sigma_{t,T}^2 = \int_t^T \nu_u^2 du.$$ 

In order to better understand this procedure, first consider an option paying $g(Y_T)B_T$ at time $T$. This has time $t$ price

$$V_t = B_t \mathbb{E}\left( \frac{g(Y_T)B_T}{B_T} \bigg| \mathcal{F}_t \right) = B_t \mathbb{E}(g(Y_T)|Y_t).$$

Comparing this to equation (3.2.7), we see that the stock pricing function is equal to $V_t$ divided by the bank account. In other words, the martingale $X_t = f(Y_t, t)$ can be written as the ratio of an option price to the numeraire under the risk-neutral measure. Consequently, we will refer to $g(y)$ as the adjusted payoff function. If this function can be analytically integrated against a normal density, then we can determine a closed form expression for $f(y, t)$. In order to satisfy assumption 3.2.2, we also require that $g(y)$ is strictly increasing and twice differentiable. In the following section we will examine a number of possible choices for $g(y)$, and the nature of the resulting LVSR models.

### 3.3 Example Models

As a first example, we will extend the shifted Black-Scholes model to account for Hull-White interest rates by letting the adjusted payoff function be $g(y) = \alpha e^y + \beta$, where $\alpha \in \mathbb{R}^+$ and $\beta \in \mathbb{R}$. Note that if $\beta = 0$ then this model reduces to the standard Black-Scholes Hull-White model, which is well studied in the literature. See for example Brigo and Mercurio (2007) for the constant parameters case. Turning back to the shifted model, equation (3.2.7) yields

$$f(y, t) = \mathbb{E}\left( \alpha e^{Y_T} + \beta | Y_t = y \right) = \alpha e^{y + \frac{1}{2} \sigma_{t,T}^2} + \beta.$$ 

Furthermore, the condition $f(0, 0) = S_0$ imposes the restriction

$$\alpha = (S_0 - \beta) e^{-\frac{1}{2} \sigma_{0,T}^2}.$$
Thus the stock price, \( S_t = D_t f(Y_t, t) \), can be written as
\[
S_t = D_t \left( (S_0 - \beta) e^{Y_t - \frac{1}{2} \sigma^2_0, t} + \beta \right),
\] (3.3.1)
and, using equation (3.2.3), the resulting volatility function is
\[
\eta(S_t, B_t, t) = \nu_t \left( 1 - \frac{D_t}{S_t} \right).
\]

Observe that that the choice of the final time horizon, \( T \), has no effect on the stock price or volatility function, and therefore does not matter.

Compared to the standard displaced diffusion model, the key difference is that the term \( D_t = A_t B_t \) is stochastic. This means that \( S_t \) is a linear combination of two correlated log-normal random variables, rather than a deterministic shift of a single log-normal random variable. Consequently, the prices of vanilla options cannot be computed using a simple modification of the Black-Scholes formula, and must instead be computed using theorem 3.2.3. We summarise this model in the following proposition.

**Proposition 3.3.1.** Under assumptions 3.2.1 to 3.2.3, let the stock pricing function be
\[
f(y, t) = (S_0 - \beta) e^{Y_t - \frac{1}{2} \sigma^2_0, t} + \beta,
\]
for some constant \( \beta \in \mathbb{R} \). Then the stock price follows the process
\[
dS_t = (r_t - y_t) S_t dt + \nu_t (S_t - \beta D_t) dW_{1,t}.
\]
We call this the “shifted exponential model”.

We can gain some insight into this model by considering the behaviour of the adjusted stock price, \( X_t = S_t D_t^{-1} \), for differing values of the displacement constant, \( \beta \). For instance, if \( \beta = 0 \), then \( \eta(S_t, B_t, t) = \nu_t \), so that equation (3.2.1) yields \( dX_t = \nu_t X_t dW_{1,t} \). Thus \( X_t \) has a log-normal distribution with time dependent volatility, \( \nu_t \). Consequently, the stock price is also log-normally distributed when \( \beta = 0 \) as it is the product of two log-normal random variables, i.e. \( S_t = A_t S_0 e^{Z_t + Y_t - \frac{1}{2} \sigma^2_0, t} \).

On the other hand, if we let \( \beta \to -\infty \), while reducing \( \nu_t \) so as to hold the “at-the-money” level of volatility constant, i.e.
\[
\eta(S_0 D_t, B_t, t) = \nu_t \left( 1 - \frac{\beta}{S_0} \right) = c_t
\]
for some time dependent value \( c_t \in \mathbb{R}^+ \), then the volatility function satisfies
\[
\eta(S_t, B_t, t) = c_t \left( \frac{S_t - \beta D_t}{S_0 - \beta} \right) \frac{S_0}{S_t} \to \frac{c_t S_0 D_t}{S_t}.
\]
Hence the model for the adjusted stock price approaches \( dX_t = c_t X_0 dW_{1,t} \) as \( \beta \to -\infty \), resulting in a normal distribution. This means that, for \( \beta \in (-\infty, 0) \), our shifted exponential model lies somewhere between the normal distribution on the one end, and the log-normal distribution on the other.

Furthermore, looking at equation (3.3.1), we see that the stock price is bounded below by \( \beta D_t \), and that negative stock prices are possible if \( \beta < 0 \). Under deterministic interest rates, Carr et al. (1999) avoid this possibility by choosing some lower barrier, \( L \), such that \( F(L, t) = 0 \), and forcing the driving process, \( Y_t \), to stay at \( L \) if it ever hits \( L \). Thus, if the stock price ever reaches 0 then it stays at 0 rather than continuing to move up and down. This works under deterministic interest rates because the PDF of a one-dimensional Brownian motion with an absorbing boundary is known analytically. However, in our case the joint PDF of \( Y_t \) and \( Z_t \) is not known when \( Y_t \) has an absorbing boundary, so we cannot adopt this approach. Instead we must either let the stock price fall below zero, or choose the function \( g \) so that \( g(y) \geq 0 \) for all \( y \in \mathbb{R} \).

Our next example is based on the hyperbolic sine model presented by Carr et al. (1999). We choose \( g(y) = \alpha \sinh(y - \beta) \), where \( \alpha \in \mathbb{R}^+ \) and \( \beta \in \mathbb{R} \). Similarly to before, we have

\[
f(y, t) = \mathbb{E} (\alpha \sinh (Y_t - \beta) | Y_t = y) = \alpha e^{\frac{1}{2} \sigma_t^2} \sinh (y - \beta),
\]

subject to the condition \( f(0, 0) = \alpha e^{\frac{1}{2} \sigma_0^2} \sinh(-\beta) = S_0 \). Thus, following assumption 3.2.2 and equation (3.2.3), the stock price and volatility function are

\[
S_t = B_t \kappa_t \sinh (Y_t - \beta), \quad (3.3.2)
\]

\[
\eta(S_t, B_t, t) = \nu_t \left(1 + \frac{\kappa_t^2 B_t^2}{S_t^2}\right)^{\frac{1}{2}} \text{sgn}(S_t), \quad (3.3.3)
\]

where \( \kappa_t \) is a deterministic function of time,

\[
\kappa_t := \frac{S_0 A_t e^{\frac{1}{2} \sigma_0^2}}{\sinh (-\beta)}. \quad (3.3.4)
\]

This lead us to the following proposition.

**Proposition 3.3.2.** Under assumptions 3.2.1 to 3.2.3, let the stock pricing function be

\[
f(y, t) = \kappa_t \sinh (y - \beta),
\]

for \( \kappa_t \) as in equation (3.3.4), and constant \( \beta \neq 0 \). Then the stock price follows the process

\[
dS_t = (r_t - y_t) S_t dt + \nu_t \left( S_t^2 + \kappa_t^2 B_t^2 \right) dW_{1,t}.
\]

We call this the “hyperbolic-sine model”.
As with the displaced exponential model, the final time horizon, \( T \), has no effect. Furthermore, looking at the behaviour of this model as \( \beta \to -\infty \), we see that \( \eta(S_t, B_t, t) \to \nu_t \text{sgn}(S_t) \), and that
\[
S_t = S_0 D_t e^{-\frac{1}{2} \sigma^2_{0,t} \sinh(Y_t - \beta) \sinh(-\beta)} \to S_0 D_t e^{Y_t - \frac{1}{2} \sigma^2_{0,t}}.
\]
Thus we have an approximately log-normal model for extreme negative values of \( \beta \). Conversely, if we let \( \beta \to 0^- \), while holding the at-the-money volatility constant, i.e.
\[
\eta(S_0 D_t, B_t, t) = \nu_t \left(1 + \frac{e^{-\sigma^2_{0,t}}}{\sinh^2(-\beta)}\right)^{\frac{1}{2}} = c_t,
\]
then the volatility function satisfies
\[
\eta(S_t, B_t, t) = c_t \left(1 + \frac{e^{-\sigma^2_{0,t}}}{\sinh^2(-\beta)}\right)^{-\frac{1}{2}} \left(1 + \frac{S_0^2 A_t^2 B_t^2 e^{-\sigma^2_{0,t}}}{\sinh^2(-\beta) S^2_t}\right)^{\frac{1}{2}} \text{sgn}(S_t) \to c_t S_0 D_t / S_t.
\]
This means that \( X_t \) approaches a normal distribution as \( \beta \to 0^- \). Thus we find that, in a similar fashion to the displaced exponential model, the hyperbolic sign model lies between the log-normal case on the one end, and the normal case on the other.

Note that the process followed by \( X_t = S_t / D_t \) under stochastic interest rates takes the same form as the process followed by \( S_t e^{-(r-y)t} \) in the hyperbolic sine model presented by Carr et al. (1999), if their absorbing barrier is set to \( L = -\infty \), where \( r \) and \( y \) are the constant interest rate and dividend rate that they adopt. Thus the behaviours of the models are quite similar, except for the stochasticity in \( B_t \), and its effects on the stock price and volatility function given in equations (3.3.2) and (3.3.3) respectively.

In the empirical tests of the next section, we will see that both models discussed above may be unable to fit the extreme negative skew in the implied volatility surface observed in real world data. Furthermore, they both allow negative stock prices, with \( S_t \) becoming negative whenever \( Y_t \) drops below \( \beta \) under the hyperbolic sine model. In order to correct these two limitations, we would like to find an adjusted payoff function that is positive, can produce extreme skews, and can also be analytically integrated against the normal density. To this end, we propose the following function which satisfies all of these conditions, unlike the stock pricing functions found in Carr et al. (1999). Let
\[
g(y) = \alpha e^{\beta y} N(\kappa y + \gamma)
\]
where \( \alpha, \beta, \kappa \in \mathbb{R}^+ \) and \( \gamma \in \mathbb{R} \). Critically, the conditional expectation \( f(y, t) = \mathbb{E}(g(Y_T) | Y_t = y) \) has
a closed form solution,

\[ f(y, t) = \alpha e^{\beta y + \frac{1}{2} \beta^2 \sigma^2_{0,T}} N \left( \frac{\kappa y + \beta \kappa \sigma^2_{I,T} + \gamma}{\sqrt{1 + \kappa^2 \sigma^2_{I,T}}} \right), \]

under the condition

\[ f(0, 0) = \alpha e^{\frac{1}{2} \beta^2 \sigma^2_{0,T}} N \left( \frac{\beta \kappa \sigma^2_{0,T} + \gamma}{\sqrt{1 + \kappa^2 \sigma^2_{0,T}}} \right) = S_0. \]

Thus, defining the constant

\[ \lambda := N \left( \frac{\beta \kappa \sigma^2_{0,T} + \gamma}{\sqrt{1 + \kappa^2 \sigma^2_{0,T}}} \right)^{-1}, \tag{3.3.5} \]

the stock price can be written as

\[ S_t = S_0 D_t \lambda e^{\beta Y_t - \frac{1}{2} \beta^2 \sigma^2_{0,t}} N \left( \frac{\kappa Y_t + \beta \kappa \sigma^2_{I,T} + \gamma}{\sqrt{1 + \kappa^2 \sigma^2_{I,T}}} \right), \tag{3.3.6} \]

and, using equation (3.2.3), the volatility function is

\[ \eta(S_t, B_t, t) = \nu_t \left( \beta + \frac{\kappa}{\sqrt{1 + \kappa^2 \sigma^2_{I,T}}} n \left( \frac{\kappa Y_t + \beta \kappa \sigma^2_{I,T} + \gamma}{\sqrt{1 + \kappa^2 \sigma^2_{I,T}}} \right) N \left( \frac{\kappa Y_t + \beta \kappa \sigma^2_{I,T} + \gamma}{\sqrt{1 + \kappa^2 \sigma^2_{I,T}}} \right)^{-1} \right), \tag{3.3.7} \]

where \( Y_t = f^{-1}(X_t, t) \), and \( n(x) \) is the standard normal PDF. These results are summarised in the proposition below.

**Proposition 3.3.3.** Under assumptions 3.2.1 to 3.2.3, let the stock pricing function be

\[ f(y, t) = S_0 \lambda e^{\beta y - \frac{1}{2} \beta^2 \sigma^2_{0,t}} N \left( \frac{\kappa y + \beta \kappa \sigma^2_{I,T} + \gamma}{\sqrt{1 + \kappa^2 \sigma^2_{I,T}}} \right) \]

for \( \lambda \) as in equation (3.3.5), and constants \( \beta, \kappa \in \mathbb{R}^+ \) and \( \gamma \in \mathbb{R} \). Then the stock price follows the process

\[ dS_t = (r_t - y_t) S_t dt + \eta(S_t, B_t, t) S_t dW_{1,t}, \]

where the volatility function \( \eta(S_t, B_t, t) \) is given in equation (3.3.7). We call this the “modified exponential model”.

Note that \( f(y, t) \) does not have an explicit inverse with respect to \( y \), so that the volatility function
3.3. Example Models

does not have a closed form expression. Nevertheless $f(y, t)$ is smooth and monotonic so its inverse can easily be computed numerically using, for example, the Newton-Raphson algorithm. It can also be cached if necessary to ensure that the volatility function can be rapidly evaluated. Moreover, the models we construct in this chapter can be simulated without using the volatility function at all. We simply simulate $Y_t$ and $Z_t$, and then use the fact that $S_t = A_t e^{Z_t f(Y_t, t)}$. Unlike a traditional LVSR model, which needs to be simulated by discretizing the underlying SDE using small time steps, our models can be exactly simulated over long time steps because $Y_t$ and $Z_t$ are jointly normal with known parameters. This is a key advantage of our approach over existing methods.

In contrast to the displaced exponential and hyperbolic sine models, the modified exponential model is sensitive to the choice of the final time horizon $T$, as is apparent in equation (3.3.6). Moreover, this model has four free parameters ($\beta$, $\kappa$, $\gamma$ and $T$) governing the stock pricing function, instead of only one. As we will see in the following section, this model is also better able to reproduce the extreme skew apparent in real market data. This, along with the fact that it avoids negative stock prices, are the primary reasons for preferring it over the first two models presented in this section.

Many other LVSR models can be produced using the above approach. We only require a suitable candidate for the adjusted payoff function $g(y)$, i.e. one that is increasing and can be analytically integrated against a normal density. For example, Carr et al. (1999) suggest using functions of the form

$$g_1(y) = \frac{P_4(e^{y-L})}{e^{n(y-L)}}, \text{ or}$$

$$g_2(y) = \frac{P_4(y-L)}{(y-L)^n},$$

where $P_4$ is a polynomial with degree $\leq 4$, and $n \leq 4$ is a non-negative integer. The degree 4 limit is imposed so that the resulting stock pricing functions can be explicitly inverted using the quartic root formula. Of course, the coefficients of these polynomials need to be carefully chosen to ensure that the function is increasing.

Although the existence of an explicit inverse is convenient when writing down the volatility function, this is not actually necessary when either calibrating or simulating the model, so we drop this condition. This means that we are free to form linear combination of previously defined adjusted payoff functions, or use polynomials up to any degree, without having to worry about analytical inversion. For example, if $f_i$ is the stock pricing function associated to the adjusted payoff function $g_i$, for $i = 1, 2, ..., m$, then we can combine the functions as follows:

$$g(y) = \sum_{i=1}^{m} \lambda_i g_i(y)$$

$$\Rightarrow f(y, t) = \sum_{i=1}^{m} \lambda_i f_i(y, t),$$

(3.3.8)
where $\lambda_i \geq 0$ so that $g(y)$ is increasing, and $\sum_{i=1}^{m} \lambda_i = 1$ so that

$$f(0, 0) = \sum_{i=1}^{m} \lambda_i S_0 = S_0.$$

As we will see in the next chapter, the LVSR models defined here can further be extended by using them as components in a mixture model. This results in a model under which vanilla option can be priced using a weighted average of their prices under each component model. This mixture-based approach has the key advantage that it extends naturally to a wide range of underlying component models, including stochastic volatility models, which is why we pursue it in addition to the methodology presented in this chapter. Furthermore, the mixture-based approach allows each component model to use a different time dependent volatility parameter, $\nu_t$, unlike the models in this section, which all allow only one time dependent parameter. Thus mixture models provide significantly more flexibility when fitting to multiple expiries.

### 3.4 Empirical Results

In this section we will present some calibration results based on real market data. We begin by fitting the Hull-White short-rate model, with piecewise constant volatility, to ICAP GBP cap volatility quotes as at 28 April 2015, obtained from Thomson Reuters Datastream. This is performed using the methodology of Brigo and Mercurio (2007), whereby the squared percentage difference between model and market cap prices is minimised.

Next, we calibrate the three models detailed in the previous section to the prices of FTSE 100 index options expiring on 15 December 2017 and 18 December 2020. We focus on such long term expiries because the impact of stochastic interest rates is more pronounced for longer term contracts. The models are calibrated by minimising the sum of squared differences between model and market call prices. In all three cases the models were calibrated within 10 seconds on an Intel i7-3820 CPU using multi-threaded code. The instantaneous correlation, $\rho_t$, between the stock price and the short-rate, is not determined by either market cap prices, or vanilla equity option prices. However, it has little effect on each model’s ability to reproduce the market’s implied volatility curve, so we set $\rho_t = 0$. In latter chapters we will be interested in the impact of stochastic interest rates on long-dated path-dependent options, which are significantly affected by correlation, and we will then adopt a range of values for $\rho_t$.

The results of our Hull-White model calibration are shown in table 3.1. We see that the one-factor Hull-White model, with piecewise constant volatility, is able to exactly fit at-the-money cap volatilities, up to the optimization algorithm’s error tolerance. Next, looking at figures 3.1, 3.2 and 3.3, we see the quality of our fits to market call option prices. Observe that the shifted exponential model and the hyperbolic sine model both produce very similar fits, and both fail to reproduce the extreme negative skew observed for the December 2017 expiry. The calibrated parameters for both
### Table 3.1: Hull-White model calibration results: cap implied volatilities

<table>
<thead>
<tr>
<th>Expiry Year</th>
<th>Model Vol</th>
<th>Market Vol</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4290</td>
<td>0.4290</td>
<td>$3.6 \times 10^{-8}$</td>
</tr>
<tr>
<td>2</td>
<td>0.6440</td>
<td>0.6440</td>
<td>$-5.8 \times 10^{-8}$</td>
</tr>
<tr>
<td>3</td>
<td>0.6955</td>
<td>0.6955</td>
<td>$3.3 \times 10^{-8}$</td>
</tr>
<tr>
<td>4</td>
<td>0.6896</td>
<td>0.6896</td>
<td>$-2.7 \times 10^{-8}$</td>
</tr>
<tr>
<td>5</td>
<td>0.6682</td>
<td>0.6682</td>
<td>$9.2 \times 10^{-9}$</td>
</tr>
<tr>
<td>6</td>
<td>0.6482</td>
<td>0.6482</td>
<td>$-4.5 \times 10^{-8}$</td>
</tr>
</tbody>
</table>

### Figure 3.1: Shifted exponential model calibration results

**Dec 2017 Expiry**

**Dec 2020 Expiry**

![Graph showing implied volatility vs strike for Dec 2017 and Dec 2020 expiries for the shifted exponential model, with market and model curves aligned.]

### Figure 3.2: Hyperbolic sine model calibration results

**Dec 2017 Expiry**

**Dec 2020 Expiry**

![Graph showing implied volatility vs strike for Dec 2017 and Dec 2020 expiries for the hyperbolic sine model, with market and model curves aligned.]


models are consistent with an approximate normal distribution for the adjusted stock price, i.e. $\beta \to -\infty$ for the former model, and $\beta \to 0^-$ for the latter. This explains why they yield such similar results.

Conversely, the modified exponential model produces a much closer fit, and is better able to match the skew for the first expiry. As explained earlier, this model also has the theoretical advantage that it avoids negative stock prices. However, it lacks an explicit inverse for the stock pricing function, and therefore lacks an explicit volatility function. Although the volatility function is not needed for calibration or simulation of the model as presented in this chapter, it is needed when simulating the mixture of component models presented in the next chapter. Thus, when simulating mixtures of this model we will use the Newton-Raphson method to numerically compute the necessary inverse function.

Note that the introduction of stochastic interest-rates does not significantly improve the fitting capability of the model, as the ability to reproduce the implied volatility curve is mostly dependent on the choice of local volatility function. However, for very long term expiries it is possible that the volatility implied by market option prices is less than the volatility caused by the stochasticity of interest rates alone. This may cause the model option prices to exceed the market prices even when stock price volatility is set to zero. However, this phenomenon was not observed in our data, so we did not have trouble calibrating the models under stochastic rates.

Naturally, it may be possible to improve the fits obtained in this section by combining multiple adjusted payoff functions using equation (3.3.8). However, we will focus on the mixture based approach due to its theoretical advantages, namely its compatibility with multiple time dependent parameters, and its extensions to stochastic volatility. Thus, the main importance of the models developed here is due to their potential use as component models in the next chapter.
3.5 Conclusion

In this chapter we have demonstrated how to construct a class of LVSR models in which the stock price can be written as the (log-normally distributed) bank account multiplied by a function of a correlated Gaussian process. This stock pricing function must satisfy a certain second order PDE, which we solve analytically using the Feynman-Kac formula.

By restricting our choice of stock pricing functions and interest rate models, so that certain expressions have analytical solutions, we are able to compute vanilla option prices using a one-dimensional integral. In this case, the calibration speed of our models may be superior to those that require iterative approximations, numerical solutions to PDEs, or Monte-Carlo simulations. However, it should be noted that we rely on a 100 point Simpson’s rule to evaluate vanilla option prices. Furthermore, our restrictions to the stock pricing function have reduced the available degrees of freedom, so that we can only provide a fit, rather than an exact match, to the implied volatility surface.

A wide range of increasing payoff functions, as long as they can be analytically integrated against a normal density, can be used to produce a valid stock pricing function. Thus the class of models that can be constructed by our method is quite large. We have given three examples, the last of which guarantees positive stock prices and is capable of producing a good fit to highly skewed implied volatility surfaces.

The models we have presented here can also be exactly simulated over long time steps without needing to evaluate their volatility functions. Thus, we do not require an explicit expression for their inverse stock pricing function, which appears in our formula for the volatility function. However, if it becomes necessary to compute this inverse, for example when constructing mixture models in the next chapter, we can easily do so numerically using, for example, the Newton-Raphson method.

Appendix 3.A Proofs

Proof of theorem 3.2.1. Define the function

\[ A(u,v) := e^{-\int_{u}^{v} a(y)dy}. \]

Solving the SDE for \( r_t \) in the risk-neutral measure yields

\[ r_t = r_0 A(0,t) + \int_{0}^{t} \theta_s A(s,t)ds + \int_{0}^{t} h(s,t)dW_{2,s}. \]
Now letting \( r_t = \tilde{r}_t + \xi_t \), where
\[
\tilde{r}_t := \int_0^t h(s, t) dW_{2,s}, \quad \xi_t := r_0 A(0, t) + \int_0^t \theta_s A(s, t) ds,
\]
we have that the log bank account, \( Z_t := \log (B_t) \), satisfies
\[
Z_t = \int_0^t r_u du = \int_0^t \int_0^u h(s, u) dW_{2,s} du + \int_0^t \xi_u du = \int_0^t H(s, t) dW_{2,s} + \int_0^t \xi_u du.
\]

Therefore the price of a zero-coupon bond expiring at time \( t \) is
\[
P_{0,t} = \mathbb{E} \left( \frac{1}{B_t} \right) = e^{\int_0^t \left( \frac{1}{2} \int_0^s H^2(s, t) ds - \int_0^s \xi_u du \right) ds}.
\]

Thus the function \( \xi_t \) is determined from the forward curve, \( f(0, t) := -\frac{d}{dt} \log (P_{0,t}) \), as follows:
\[
\log (P_{0,t}) = \int_0^t \left( \frac{1}{2} H^2(s, t) - \xi_s \right) ds,
\]
\[
\Longrightarrow \frac{d}{dt} \log (P_{0,t}) = -\xi_t + \frac{1}{2} \int_0^t \frac{\partial}{\partial t} H^2(s, t) ds,
\]
\[
\Longrightarrow \xi_t = f(0, t) + \int_0^t h(s, t) H(s, t) ds.
\]

Now changing to the \( T \)-forward measure, as explained in section 12.1.1 of Brigo and Mercurio (2007), we have that
\[
dW_{2,t}^T = dW_{2,t} + H(t, T) dt
\] (3.A.1)
is a simple Brownian motion in that measure. Thus the SDE for \( r_t \) in the \( T \)-forward measure is
\[
dr_t = (\theta_t - \psi_t H(t, T) + a_t r_t) dt + \psi_t dW_{2,t}^T,
\]
which on solving yields

\[ r_t = r_0 A(0, t) + \int_0^t (\theta_s - \psi_s H(s, T)) A(s, t) \, ds + \int_0^t h(s, t) dW_{2,s}^T \]

\[ = \xi_t - \int_0^t h(s, t) H(s, T) \, ds + \int_0^t h(s, t) dW_{2,s}^T \]

\[ = f(0, t) - \int_0^t h(s, t) (H(s, T) - H(s, t)) \, ds + \int_0^t h(s, t) dW_{2,s}^T \]

\[ = f(0, t) - \int_0^T \int h(s, t) h(s, z) \, dz \, ds + \int_0^t h(s, t) dW_{2,s}^T, \quad (3.A.2) \]

Therefore the log bank account at expiry is

\[ Z_T = \int_0^T f(0, u) \, du - \int_0^T \int \int h(s, u) h(s, z) \, dz \, ds \, du + \int_0^T \int h(s, u) dW_{2,s}^T \, du \]

\[ = - \log (P_{0,T}) - \int_0^T \int \int h(s, u) h(s, z) \, dz \, ds \, du + \int_0^T H(s, T) dW_{2,s}^T \]

\[ = - \log (P_{0,T}) - \frac{1}{2} \int_0^T H^2(s, T) \, ds + \int_0^T H(s, T) dW_{2,s}^T, \quad (3.A.3) \]

where the last equality is due to the fact that the integral of \( h(s, u) h(s, z) \) over the triangle \( \{(z, u) : z \in [u, T], u \in [s, T]\} \) is equal to half the integral of that term over the square \( \{(z, u) : (z, u) \in [s, T] \times [s, T]\} \). Finally, comparing equations (3.A.2) and (3.A.3), we see that \((r_T, Z_T)\) has a bivariate normal distribution in the \( T\)-forward measure, with the parameters given in theorem 3.2.1. \( \square \)

**Proof of theorem 3.2.2.** From assumption 3.2.2 we have that

\[ dY_t = \nu_t dW_{1,t}. \]

Changing to the \( T\)-forward measure using equation (3.A.1) yields

\[ dW_{1,t} = \rho_t dW_{2,t} + \sqrt{1 - \rho_t^2} d\tilde{W}_{1,t} \]

\[ = -\rho_t H(t, T) dt + \rho_t dW_{2,t}^T + \sqrt{1 - \rho_t^2} d\tilde{W}_{1,t}, \]
where $W^T_{2,t}$ and $\tilde{W}_{1,t}$ are independent standard Brownian motions in that measure. Thus

$$Y_t = -\int_0^t \nu_u \rho_u H(u, T) du + \int_0^t \nu_u \rho_u dW^T_{2,u} + \int_0^t \nu_u \sqrt{1 - \rho_u^2} d\tilde{W}_{1,u}. \tag{3.A.4}$$

Finally, comparing equations (3.A.2), (3.A.3) and (3.A.4), we see that $(Y_T, r_T, Z_T)$ has a multivariate normal distribution in the $T$-forward measure, with the parameters given in theorem 3.2.2. □
Chapter 4

Local Volatility under Stochastic Interest Rates Using Mixture models

4.1 Introduction

In this chapter we show how to use a mixture model to greatly simplify the calibration of a local volatility function under stochastic interest rates. Unlike the LVSR models presented in the previous chapter, these mixture models may have an arbitrary number of time-dependent volatility parameters, and are therefore better able to fit implied volatility smiles at multiple expiries. They can also be extended to more general multi-factor models, such as those including both stochastic volatility and stochastic interest rates.

A Markov functional approach, such as that developed by Fries and Eckstaedt (2011), may be considered more general than the mixture models presented in this chapter. This is because it is able to exactly match market option prices by allowing the functional form to vary with time in a piece-wise constant fashion. However, in order to maintain analytical tractability, the Markov functional models presented in the previous chapter do not have the ability to select an arbitrary piece-wise constant functional form. Thus, the mixture model framework developed in this chapter is necessary to regain the flexibility to match option prices at multiple expiries.

Our approach is quite fast compared to previous methods because vanilla options are simply priced using the weighted average of each component model's price. We present two different methods for calculating the local volatility function. The first approximates the joint distribution of the stock price and short rate using a mixture of component joint distributions. This yields a straightforward analytical formula for volatility as a function of the stock price and time.
The second method exactly reproduces the joint distribution, in the forward measure, of a given mixture model by allowing the volatility function to depend on all the included state variables, rather than just the stock price and time. This extended volatility function is determined by equating the PDE governing the joint distribution in the main model to the weighted sum of the PDEs governing the joint distribution in the component models. The generality of this approach means that it can be used to combine many different types of multivariate models, with three examples given in this chapter.

We begin, in section 4.2, by using an approximation for the joint distribution of the short rate and stock price to recover an analytical formula for the local volatility function. Next, in section 4.3, we consider a general multivariate diffusion, and determine the drift and volatility that allow it to match the joint density in the $t$-forward measure of a given mixture model. This approach is used in section 4.4 to construct an example mixture model in which volatility is a function of the stock price, short rate, bank account and time. Next, in section 4.5, we look at the accuracy of our approximate method, and highlight the potential impact of stochastic rates when pricing long term path dependent options, specifically up-and-out call options. Finally, we conclude in section 4.6.

### 4.2 Approximate Local Volatility Using Mixture Models

Consider a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ satisfying the usual conditions. The following assumption describes a local volatility model that allows for correlated stochastic interest rates.

**Assumption 4.2.1.** We assume that markets are frictionless and that there exists an equivalent risk-neutral measure, $\mathbb{Q}$, such that the stock price, $S_t$, and the short rate, $r_t$, follow diffusive processes of the form

\[
\begin{align*}
    dS_t &= (r_t - y_t) S_t \, dt + \eta(S_t, t) S_t \, dW_{1,t}, \\
    dr_t &= \mu(r_t, t) \, dt + \sigma(r_t, t) \, dW_{2,t},
\end{align*}
\]

where $y_t$ is the deterministic dividend rate, $\eta$ is the local volatility function, $\mu$ is the short rate’s drift, $\sigma$ is the short rate’s volatility, and $(W_1, W_2)$ is a two-dimensional $\mathcal{F}_t$-adapted Brownian motion with instantaneous correlation $\rho_t$.

We focus on a one-factor short-rate model for simplicity, but the extension to multi-factor models is straightforward. Given a complete surface of market call option prices for all strikes and expiries, it can be shown (Benhamou et al., 2008) that the local volatility function, $\eta(K, T)$, satisfies

\[
\eta^2(K, T) = \frac{\partial}{\partial K} C - KP_{0,T} \mathbb{E}^T \left( r_T 1_{\{S_T > K\}} \right) + y_t \left( C - K \frac{\partial}{\partial K} C \right),
\]

where $C = C(K, T)$ is the time-zero price of a call option with strike $K$ that expires at time $T$, $P_{0,T}$ is the time-zero price of a zero-coupon bond that pays $1$ at time $T$, and $\mathbb{E}^T(\cdot)$ is the expectation in
the $T$-forward measure. All the terms on the right hand side of equation (4.2.2) are known except for $\mathbb{E}^T (r_T \mathbb{I}_{S_T > K})$. However, we can approximate this unknown term for all values of $K$ and $T$ by first estimating the joint density of $S_T$ and $r_T$ using a mixture model. This approach also allows us to interpolate between the vanilla option prices actually available in the market. The details of our mixture model are given in the assumption below.

**Assumption 4.2.2.** Let $\{S_{k,t} : k = 1, 2, ..., n\}$ be a set of $n$ component diffusions, each with their own deterministic dividend rate, $y_{k,t}$, and local volatility function, $\eta_k$, which both depend on some parameter set $\theta_k$, such that

$$dS_{k,t} = \left(r_t - y_{k,t}(\theta_k)\right) S_{k,t} dt + \eta_k (S_{k,t}, t, \theta_k) S_{k,t} dW_{1,t}. \quad (4.2.3)$$

All the component diffusions share the same short rate process, $r_t$, and pair of driving Brownian motions, $(W_1, W_2)$, as in assumption 4.2.1. Associated to each component diffusion is a non-negative mixture weight, $\lambda_k \geq 0$, such that $\sum_{k=1}^{n} \lambda_k = 1$. In the $t$-forward measure, we assume that the marginal density of $S_t$ equals the weighted sum of the marginal densities of $S_{k,t}$ in that measure, i.e.

$$\phi_{S_{k,t}} (x) = \sum_{k=1}^{n} \lambda_k \phi_{S_{k,t}} (x). \quad (4.2.4)$$

Note that the constants $\lambda_k$ are not allowed to depend on time or the stock price. Therefore the existence of a model for $S_t$ satisfying equation (4.2.4) in the $t$-forward measure is not guaranteed, and we take this to be part of the assumption.

In order to maintain tractability, the component local volatilities, $\eta_k$, would in practice be functions of $t$ only, but we present the general case here. We also allow each component diffusion to have a different dividend rate, $y_{k,t}$, to ensure that the model has the flexibility to fit a skewed market smile. However, as explained by Brigo et al. (2003), $y_{k,t}$ must be chosen to satisfy the forward pricing equation,

$$S_0 e^{-\int_0^T y_s ds} = P_{0,T} \mathbb{E}^T (S_T) = S_0 \sum_{k=1}^{n} \lambda_k e^{-\int_0^T y_{k,t} dt}, \quad \forall T \geq 0 \quad (4.2.5)$$

$$\Rightarrow \sum_{k=1}^{n} \lambda_k (y_{k,T} - y_T) e^{-\int_0^T y_{k,t} dt} = 0, \quad \forall T \geq 0.$$

Assuming that the short-rate model has been fitted to interest rate market data, and that an analytical formula for vanilla option prices is available for each component diffusion, we can fit the mixture model to observed market call prices by minimizing, for example, the following objective
function.

\[
O (\lambda_1, \lambda_2, ..., \lambda_n, \theta_1, \theta_2, ..., \theta_n) = \sum_{i=1}^{n_{\text{obs}}} \left( C (K_i, T_i) - \sum_{k=1}^{n} \lambda_k C_k (K_i, T_i, \theta_k) \right)^2. \tag{4.2.6}
\]

Here \(n_{\text{obs}}\) is the number of call option prices observable in the market, \(C (K_i, T_i)\) in the actual price of call option number \(i\), and \(\sum_{k=1}^{n} \lambda_k C_k (K_i, T_i, \theta_k)\) is its model price.

Note that, because any expectation of the form \(\mathbb{E}_T (f (S_T))\) in a mixture model is equal to the weighted average of the expectations in the component models, the price of a call option is simply the weighted average of the component models’ prices. Also, as option prices do not contain much information regarding \(\rho_t\), we must estimate it first (using time-series data for example) and then use our objective function to determine the remaining parameters, holding \(\rho_t\) fixed.

The above procedure must be conducted subject to the constraint on \(y_{k,t}\) given by equation (4.2.5). However, in the case of piece-wise constant \(y_{k,t}\), it is possible to re-parametrize the objective function in such a way that the minimization can be performed in an unconstrained manner. We refer the reader to Rebonato and Cardoso (2004) for further details.

After fitting a mixture model to market prices, our task is to find the volatility function, \(\eta (K, T)\), such that the prices of vanilla options are the same in the local volatility model and the mixture model. We would like to apply equation (4.2.2), but to do so we need to know the joint distribution of \(S_T\) and \(r_T\) in the \(T\)-forward measure. However, this is not easy to compute given only the marginal distribution of \(S_T\), as described by equation (4.2.4), and a correlated short-rate process, as described by equation (4.2.1). Therefore we proceed by making the following approximation.

**Approximation 4.2.1.** We assume that the joint distribution of the short rate and stock price is approximately equal to the mixture of the joint distributions in each component model, i.e.

\[
\phi_{S_t, r_t} (x_1, x_2) \approx \sum_{k=1}^{n} \lambda_k \phi_{S_k, r_t} (x_1, x_2).
\]

Thus

\[
\mathbb{E}_T (r_T \mathbb{1}_{\{S_T > K\}}) \approx \sum_{k=1}^{n} \lambda_k \mathbb{E}_T^k (r_T \mathbb{1}_{\{S_T > K\}}).
\]

Next, by applying equation (4.2.2) to each component model and rearranging, we get

\[
KP_{0,T} \mathbb{E}_T (r_T \mathbb{1}_{\{S_T > K\}}) = \frac{\partial}{\partial T} C_k - \frac{1}{2} K^2 \eta_k^2 (K, T) \frac{\partial^2}{\partial K^2} C_k + y_{k,t} \left( C_k - K \frac{\partial}{\partial K} C_k \right)
\]

\[
\implies KP_{0,T} \mathbb{E}_T (r_T \mathbb{1}_{\{S_T > K\}}) \approx \sum_{k=1}^{n} \lambda_k \left[ \frac{\partial}{\partial T} C_k - \frac{1}{2} K^2 \eta_k^2 (K, T) \frac{\partial^2}{\partial K^2} C_k + y_{k,t} \left( C_k - K \frac{\partial}{\partial K} C_k \right) \right].
\]
4.3. Multivariate Local Volatility Using Mixture Models

Feeding this back into equation (4.2.2) yields the desired approximation for $\eta^2 (K, T)$.

**Proposition 4.2.1.** Under approximation 4.2.1, and assumptions 4.2.1 and 4.2.2, the square of the local volatility function is approximately

$$
\eta^2 (K, T) \approx \sum_{k=1}^{n} \lambda_k \left[ K^2 \eta_k^2 (K, T) \frac{\partial^2}{\partial K^2} C_k - 2 y_{k,t} \left( C_k - K \frac{\partial}{\partial K} C_k \right) \right] + 2 y_t \left( C - K \frac{\partial}{\partial K} C \right) K^2 \frac{\partial^2}{\partial K^2} C
$$

We see that the squared local volatility function is the weighted average of the component squared local volatility functions, plus a term due to the differing drifts in each component diffusion. Note that the weights in the first term can be written as

$$
w_k (K, T) = \lambda_k \frac{\partial^2}{\partial K^2} C_k
$$

This approach provides an analytical approximation for the local volatility function, under correlated stochastic interest rates, based on a given mixture model. Alternatively, if we wish to model the joint distribution of the stock price and short-rate exactly, or want to mix different types of component models, we can use the general multivariate approach detailed below.

**4.3 Multivariate Local Volatility Using Mixture Models**

Our goal in this section is to find a multivariate diffusion process whose joint density, in the forward measure, exactly matches that of a given mixture model. Unlike the previous section, we make no approximations. However, in order to match the joint density of each process included in the model, we allow the local volatility function to depend on all the processes, rather than just the stock price and time. Although this approach is non-standard, it gives us the added freedom necessary to fit the joint distribution.

We begin by specifying a general multivariate diffusion model in the risk-neutral measure. This model could include, for example, one process for the stock price, one for the short-rate, and one for the bank account.

**Assumption 4.3.1.** We assume that markets are frictionless and that there exists an equivalent risk-neutral measure, $Q$, such that the vector of economic variables of interest, $X_t \in \mathbb{R}^m$, follows a diffusive
process of the form

\[ dX_t = \mu(X_t, t) \, dt + \Sigma(X_t, t) \, dW_t, \]

where \( W_t \) is a standard \( m \)-dimensional \( \mathcal{F}_t \)-adapted \( \mathbb{Q} \)-Brownian motion. The drift coefficient, \( \mu : \mathbb{R}^m \times [0, \infty) \to \mathbb{R}^m \), and diffusion coefficient, \( \Sigma : \mathbb{R}^m \times [0, \infty) \to \mathbb{R}^{m \times m} \), must be measurable functions such that \( X_t \) has a unique strong solution. Let \( \phi_t(x) \) be the joint density of \( X_t \) in the \( t \)-forward measure. We will use \( X_{i,t}, \mu_i, \sigma_{i,j} \) and \( W_{i,t} \) to refer to the elements of \( X_t \), \( \mu \), \( \Sigma \) and \( W_t \) respectively.

Next, we define our mixture model in terms of \( n \) component diffusions of the same form as in assumption 4.3.1. Note that all of these component diffusions share the same initial value, \( X_0 \), so that the resulting mixture diffusion has a deterministic starting point.

**Assumption 4.3.2.** Let \( \{X_t^{(k)} : k = 1, 2, \ldots, n\} \) be a set of \( n \) component diffusions of the form

\[ dX_t^{(k)} = \mu^{(k)}(X_t, t) \, dt + \Sigma^{(k)}(X_t, t) \, dW_t, \]

where the drift coefficients, \( \mu^{(k)} : \mathbb{R}^m \times [0, \infty) \to \mathbb{R}^m \), and diffusion coefficients, \( \Sigma^{(k)} : \mathbb{R}^m \times [0, \infty) \to \mathbb{R}^{m \times m} \), must be measurable functions such that \( X_t^{(k)} \) has a unique strong solution. All the component diffusions share the same driving Brownian motion, \( W_t \), and starting point, \( X_0^{(k)} = X_0 \), as in assumption 4.3.1. Associated to each component diffusion is a non-negative mixture weight, \( \lambda_k \geq 0 \), such that \( \sum_{k=1}^n \lambda_k = 1 \). Let \( \phi_{k,t}(x) \) be the joint density of \( X_t^{(k)} \) in the \( t \)-forward measure. We will use \( X_{i,t}^{(k)}, \mu_i^{(k)} \) and \( \sigma_{i,j}^{(k)} \) to refer to the elements of \( X_t^{(k)} \), \( \mu^{(k)} \) and \( \Sigma^{(k)} \) respectively.

Now our goal is to determine the functions \( \mu \) and \( \Sigma \) in assumption 4.3.1 such that the joint density of the state variables in the \( t \)-forward measure is equal to the weighted average of the joint densities implied by each component diffusion, i.e.

\[ \phi_t(x) = \sum_{k=1}^n \lambda_k \phi_{k,t}(x), \quad x \in \mathbb{R}^m. \]  

(4.3.1)

Such a model is desirable because it allows vanilla options to be priced rapidly using the weighted average of each component model’s price. For example, a call option with expiry \( T \) and strike \( K \) has price

\[ C(K, T) = \sum_{k=1}^n \lambda_k C_k(K, T), \]  

(4.3.2)

where \( C_k(K, T) \) is the price of that call option under component model \( k \). The fundamental equation we will use to achieve this goal is the PDE governing the evolution of the \( t \)-forward joint density through time, which is given in the following theorem from Overhaus et al. (2007).
Theorem 4.3.1. (Overhaus et al., 2007) Under assumption 4.3.1, suppose that there exists a measurable function, $g : \mathbb{R}^m \times [0, \infty) \to \mathbb{R}$, such that $g(\mathbf{X}_t, t) = r_t$ is the short rate for all $t \geq 0$. Let $f(0, t) := -\frac{\partial}{\partial t} \log P(0, t)$ be the instantaneous forward rate at time 0 for maturity $t$. Then the joint density of $\mathbf{X}_t$ in the $t$-forward measure, $\phi_t(\mathbf{x})$, satisfies

$$
\frac{\partial}{\partial t} \phi_t = (f(0, t) - g(\mathbf{x}, t)) \phi_t - \sum_{i=1}^{m} \frac{\partial}{\partial x_i} (\mu_i \phi_t) + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2}{\partial x_i \partial x_j} (D_{i,j} \phi_t),
$$

(4.3.3)

where $D_{i,j} = \sum_{l=1}^{m} \sigma_{i,l} \sigma_{j,l}$, and $\mathbf{x} := (x_1, x_2, ..., x_m)$. Note that we have omitted the $(\mathbf{x}, t)$ dependence in $\mu_i$, $\sigma_{i,j}$ and $D_{i,j}$, and also the $\mathbf{x}$ dependence in $\phi_t$, in order to ease the notation.

Although equation (4.3.3) is similar to the Fokker-Plank equation, it differs in that the measure varies with $t$ instead of remaining fixed. Thus it includes the additional term $(f(0, t) - g(\mathbf{x}, t)) \phi_t$.

Now differentiating equation (4.3.1) with respect to $t$, we have

$$
\frac{\partial}{\partial t} \phi_t = \sum_{k=1}^{n} \lambda_k \frac{\partial}{\partial t} \phi_{k,t}.
$$

Applying theorem 4.3.1 then yields

$$
(f(0, t) - g(\mathbf{x}, t)) \phi_t - \sum_{i=1}^{m} \frac{\partial}{\partial x_i} (\mu_i \phi_t) + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2}{\partial x_i \partial x_j} (D_{i,j} \phi_t)
$$

(4.3.4)

$$
= \sum_{k=1}^{n} \lambda_k \left( (f_k(0, t) - g_k(\mathbf{x}, t)) \phi_{k,t} - \sum_{i=1}^{m} \frac{\partial}{\partial x_i} (\mu_i^{(k)} \phi_{k,t}) + \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2}{\partial x_i \partial x_j} (D_{i,j}^{(k)} \phi_{k,t}) \right),
$$

where the short-rate function and forward curve in each component model are given by $g_k$ and $f_k$ respectively.

To simplify equation (4.3.4) we will assume that all the component models share the same $g_k$ and $f_k$ as the main model. This means that every model’s short-rate process is the same function of the state variables, and is calibrated to the same yield curve. However, any parameters not affecting the yield curve calibration are allowed to vary between the models. These conditions are summarized below.

Assumption 4.3.3. Under assumptions 4.3.1 and 4.3.2, suppose that there exists a measurable function, $g : \mathbb{R}^m \times [0, \infty) \to \mathbb{R}$, such that $g(\mathbf{X}_t, t) = r_t$ is the short rate for all $t \geq 0$ in the main model. Assume that the short-rate in every component model is also given by this same function, i.e.

$$
r_t^{(k)} = g(X_t^{(k)}, t), \quad k = 1, 2, ..., m.
$$
Also assume that every component model is calibrated to the market forward curve i.e.

\[ f_k(0, t) = f_{\text{mkt}}(0, t), \quad k = 1, 2, ..., m. \]

Typically the main model and all the component models would explicitly include the short-rate as the \( i \)th element of the state vector, for some integer \( i \). In this case, the function \( g \) is the projection onto the \( i \)th coordinate of \( x \), i.e. \( g(x, t) = x_i \). Note that equation (4.3.1) and assumption 4.3.3 together imply that the main model is also calibrated to the market forward curve. This is because the instantaneous forward rate is equal to the expected value of the short rate in the \( t \)-forward measure, i.e.

\[ f(0, t) = \mathbb{E}^t(r_t) \]

\[ = \int_{\mathbb{R}^m} g(x, t) \phi_t(x) dx \]

\[ = \int_{\mathbb{R}^m} g(x, t) \sum_{k=1}^{n} \lambda_k \phi_{k,t}(x) dx \]

\[ = \sum_{k=1}^{n} \lambda_k \mathbb{E}^t(r_t^{(k)}) \]

\[ = f_{\text{mkt}}(0, t) \]

Thus, equation (4.3.4) reduces to

\[ \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \left( \mu_i \phi_t \right) - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2}{\partial x_i \partial x_j} \left( D_{i,j} \phi_t \right) \]

\[ = \sum_{i=1}^{m} \frac{\partial}{\partial x_i} \left( \sum_{k=1}^{n} \lambda_k \mu_i^{(k)} \phi_{k,t} \right) - \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \frac{\partial^2}{\partial x_i \partial x_j} \left( \sum_{k=1}^{n} \lambda_k D_{i,j}^{(k)} \phi_{k,t} \right). \]

A straightforward way to solve this equation is to equate each term in brackets on the left hand side to the corresponding term on the right hand side. This leads to the following theorem, which is the main result of this chapter.

**Theorem 4.3.2.** Under assumptions 4.3.1 to 4.3.3, suppose that the drift coefficient, \( \mu \), and diffusion coefficient, \( \Sigma \), in the main model satisfy

\[ \mu = \frac{\sum_{k=1}^{n} \lambda_k \mu^{(k)} \phi_{k,t}}{\sum_{k=1}^{n} \lambda_k \phi_{k,t}}, \quad \Sigma' = \frac{\sum_{k=1}^{n} \lambda_k \Sigma^{(k)} \Sigma^{(k)'} \phi_{k,t}}{\sum_{k=1}^{n} \lambda_k \phi_{k,t}}. \]  

(4.3.5)

Also suppose that the resulting SDE has a unique strong solution whose joint density in the \( t \)-forward measure, \( \phi_t \), is the unique solution to equation (4.3.3). Then \( \phi_t \) is equal to the weighted
average of the component joint densities, i.e

$$\phi_t = \sum_{k=1}^{n} \lambda_k \phi_{k,t}. $$

Note that the right hand side of equation (4.3.5) is a positive linear combination of positive semidefinite matrices, so is itself positive semidefinite. Therefore there exists at least one solution, $\Sigma$, to this equation. The example later in this chapter will use the Cholesky decomposition.

Using theorem 4.3.2 we can write down the drift and volatility of a process whose joint distribution in the $t$-forward measure is equal to that of a given mixture model. However, for this to be a valid arbitrage free model, we must ensure that the drift of the stock price is equal to the short-rate minus the dividend rate all multiplied by the stock price. For example, if $X_t^{(k)}$ represent the stock price in component model $k$, then we will set

$$\mu_t^{(k)} = (g(x, t) - y_t) x_l,$$

for $k = 1, 2, ..., m$, so that the drift of the stock price is

$$\mu_t = \frac{\sum_{k=1}^{n} \lambda_k \mu_{t}^{(k)} \phi_{k,t}}{\sum_{k=1}^{n} \lambda_k \phi_{k,t}} = (g(x, t) - y_t) x_l.$$

Furthermore, we must be careful when pricing derivatives whose payoffs depend on anything other than the state variables at some fixed expiry. For instance, a swaption’s payoff at expiry depends on multiple points of the yield curve (rather than just the short rate). Therefore its price may not equal the weighted average of the prices implied by each component model’s short-rate process.

However, we can avoid this problem by making sure that the short-rate follows the same process in every component model, because this means that the short-rate in the main model will also follow this same process. Thus the prices of all interest rate derivatives will be the same in all the component models and the main model. Note that we may still allow the instantaneous correlation between the short-rate and stock price to vary between the component models while holding the short-rate process fixed.

### 4.4 Example Mixture Model

In this section we will produce an example mixture model using component models taken from the previous chapter, namely the shifted exponential model, the hyperbolic-sine model, and the modified exponential model. Each of these three models essentially combines Hull-White interest rates with a particular parametric form for the local volatility function. The setup of our mixture
model is summarised below.

**Assumption 4.4.1.** Under assumption 4.3.2, let \( X_t^{(k)} := (S_{k,t}, r_{k,t}, B_{k,t}) \), be the stock price, short-rate and bank account in component model \( k \), such that

\[
dS_{k,t} = (r_{k,t} - y_t) S_{k,t} dt + \chi_k (S_{k,t}, B_{k,t}, t) dW_{1,t},
\]
\[
dr_{k,t} = (\theta_t - a_t r_{k,t}) dt + \psi_t \left( \rho_{k,t} dW_{1,t} + \sqrt{1 - \rho_{k,t}^2} dW_{2,t} \right),
\]
\[
 dB_{k,t} = r_t B_{k,t} dt,
\]

where \( \theta_t, a_t, \psi_t \) and \( \rho_{k,t} \) are deterministic functions of \( t \), with \( \rho_{k,t} \in (-1, 1) \), and the diffusion coefficient, \( \chi_k (x_1, x_3, t) \), depends on the choice of component model \( k \) as follows:

\[
\chi_k (x_1, x_3, t) = \begin{cases} 
\nu_k (x_1 - \beta_k A_t x_3), & \text{for the shifted exponential model} \\
\nu_k \sqrt{x_1^2 + \kappa_{k,t}^2 x_3^2}, & \text{for the hyperbolic-sine model} \\
\eta_k (x_1, x_3, t) x_1, & \text{for the modified exponential model}
\end{cases}
\]

The volatility function, \( \eta_k (x_1, x_3, t) \), for the modified exponential model is given in equation (3.3.7). The parameters governing each type of component model are as described in propositions 3.3.1 to 3.3.3. All of these parameters, except for \( y_t, \theta_t, a_t \) and \( \psi_t \), may vary with \( k \).

In order to calibrate our model, we begin by fitting the short-rate parameters, \( \theta_t, a_t \) and \( \psi_t \), to market cap prices using the formulas in Hull and White (1990). As explained previously, these parameters are used in every component model so that the short-rate process in the final mixture model will be unchanged. Next, we fix a value for correlation, \( \rho_{k,t} \), and calibrate the remaining parameters to the market prices of call options by minimizing the objective function given in equation (4.2.6). This function can be computed rapidly (because it only involves a one-dimensional numerical integral) by using equation (3.2.5) to price call options in each component model. This is possible because the stock pricing function and interest rate model have been restricted in such a way to allow for analytical solutions to certain expressions. For more general models a higher-dimensional numerical integration would be required. Note that \( \rho_{k,t} \) is allowed to vary with time and between the component models. However, in practice a single fixed value may be estimated using time series data.

We can now use theorem 4.3.2 to determine the diffusion coefficient, \( \Sigma (x, t) \), that is consistent with this mixture model. Specifically, we equate \( \Sigma (x, t) \) to the Cholesky decomposition of the instantaneous covariance matrix defined in equation (4.3.5). The resulting formula for \( \Sigma (x, t) \) is fully state dependent because it depends on the joint densities of \( X_t^{(k)} \) for each component model. From chapter 3, we know that \( r_{k,t}, \log (B_{k,t}) \) and

\[
Y_{k,t} = f_k^{-1} \left( \frac{S_{k,t}}{A_t B_{k,t}}, t \right)
\]
are jointly normal in the $t$-forward measure, with the parameters given in theorems 3.2.1 and 3.2.2. Therefore, the necessary joint densities can be found using a straightforward transformation of the joint normal density. Our final model is summarised in the following proposition.

**Proposition 4.4.1.** Under assumptions 4.3.1 and 4.4.1, let $X_t := (S_t, r_t, B_t)$, be the stock price, short-rate and bank account. Also let

$$dS_t = (r_t - y_t) S_t dt + \sqrt{\sum_{k=1}^{n} \lambda_k \chi_k^2 (S_t, B_t, t) \phi_{k,t} (X_t)} dW_{1,t},$$

(4.4.1)

where $r_t$ follows the Hull-White model with state-dependent correlation coefficient

$$\rho (X_t, t) = \frac{\sum_{k=1}^{n} \lambda_k \chi_k (S_t, B_t, t) \rho_{k,t} \phi_{k,t} (X_t)}{\sqrt{\left( \sum_{k=1}^{n} \lambda_k \phi_{k,t} (X_t) \right) \left( \sum_{k=1}^{n} \lambda_k \chi_k^2 (S_t, B_t, t) \phi_{k,t} (X_t) \right)}}.$$  

(4.4.2)

Then, under the assumptions of theorem 4.3.2,

$$\phi_t (x) = \sum_{k=1}^{n} \lambda_k \phi_{k,t} (x).$$

Note that the standard Black-Scholes Hull-White (BSHW) model is a special case of the shifted exponential model with the shifting constant set to zero. Under this model the stock price is log-normally distributed in the forward-measure, so vanilla options are priced using a standard Black-Scholes type formula. Furthermore, the bank account does not appear in the SDEs for $S_t$ or $r_t$. Therefore a mixture consisting purely of BSHW components needs only two state variables, rather than three. However, as shown by Brigo and Mercurio (2001), a mixture of log-normal distributions with identical drifts cannot produce a volatility skew, since the minimum of the implied volatility curve will always occur at the forward price, regardless of the parameters. Therefore, when modelling an option market that displays a significant skew, a mixture of BSHW models is insufficient. This is the key reason why the the parametric LVSR models developed in chapter 3 are necessary when constructing a LVSR mixture model.

### 4.5 Empirical Results

To highlight the potential impact of interest rate stochasticity on the prices of exotic options, and also test the accuracy of the approximate method detailed in section 4.2, we perform an example calibration to FTSE 100 index option data. This is the same data as was used in section 3.4, so our results are directly comparable. Furthermore, as we are using the same GBP cap volatility quotes
as before, our Hull-White model calibration is unchanged.

Taking the interest rate process as given, we fit three forms of mixture models to market call and put options expiring on 15 December 2017 and 18 December 2020. The first form consists of four Black-Scholes Hull-White type components, as described by assumption 4.2.2, each with dividend rate \( y_{k,t} \) and piecewise constant volatility \( \eta_{k,t} \). The different dividend rates allow each component stock price process to have a different drift in the risk neutral measure, and gives the model the flexibility to fit a wide range of skews. Note that the time dependent dividend rate for the last component is determined from the piecewise constant dividend rates for the first three components using equation (4.2.5).

The second form of model consists of four shifted exponential component models, as described in assumption 4.4.1, each with shifting parameter \( \beta_k \), and piecewise constant volatility \( \nu_{k,t} \). The volatility function and instantaneous correlation necessary to simulate this mixture model are computed according to equations (4.4.1) and (4.4.2) respectively. Note that all the component models have the same dividend rate, and therefore an identical drift in the risk neutral measure, as this is necessary to apply the results of section 4.4. This means that each component stock price process must match the market forward price at each expiry, unlike the mixture of BSHW models with different drifts, which can have a different forward price for each component. Thus all skew is provided by the shifting constants, \( \beta_k \).

The third form of model is similar to the second, except we use four modified exponential component models, each with parameters \( \beta_k, \kappa_k, \gamma_k \) and \( \nu_{k,t} \). The details of this type of component model can be found in proposition 3.3.3. The reason we include this example is because it is better able to reproduce the extreme skew present in the implied volatility curve for the 15 December 2017 expiry.

The instantaneous correlation parameter, \( \rho \), between equity prices and interest rates is not implied by either market cap prices, or vanilla equity option prices. Therefore, following the approach of Hörfelt and Torné (2010), we perform three separate calibrations for \( \rho \) equal to 0.4, 0 and −0.4, as historical correlation is typically within this range.

The quality of our fits are shown in figures 4.1, 4.2 and 4.3. Looking at figure 4.1, we see that the mixture of BSHW models with different drifts is able to closely match market implied volatilities, using only four component models. Furthermore, by increasing the number of components used, it is possible to achieve an even better fit. Note that we have only shown results for \( \rho = 0.4 \), as the fits for other values of \( \rho \) were very similar.

However, looking at figure 4.2, we see that the mixture of shifted exponential models with identical drifts has some trouble matching the market prices of call options for the first expiry. This is because the underlying components are unable to reproduce this expiry’s extreme skew, and we are not allowed to give each component a different drift (as can be done for the first form of mixture model). Hence, even if the second expiry is removed from the calibration, the fit for the first expiry cannot be improved significantly. Likewise, a mixture of hyperbolic-sine component models is unable to match the observed skew for the first expiry, and yields a fit almost identical to the shifted
4.5. Empirical Results

Figure 4.1: BSHW mixture model with different drifts: calibration for $\rho = 0.4$

![Plot of implied volatility vs strike for Dec 2017 Expiry and Dec 2020 Expiry for Model and Market]

Figure 4.2: Shifted exponential mixture model with identical drifts: calibration for $\rho = 0.4$

![Plot of implied volatility vs strike for Dec 2017 Expiry and Dec 2020 Expiry for Model and Market]

Figure 4.3: Modified exponential mixture model with identical drifts: calibration for $\rho = 0.4$

![Plot of implied volatility vs strike for Dec 2017 Expiry and Dec 2020 Expiry for Model and Market]
exponential model. Therefore we have not given results for this type of mixture model.

On the other hand, looking at figure 4.3, we see that a mixture of modified exponential models produces a much better fit to market option prices for the first expiry. Furthermore, unlike the hyperbolic-sine and shifted exponential models, it guarantees positive stock prices. The key drawback of the modified exponential model is that its volatility function does not have a closed form expression, and its evaluation requires the numerical inversion of the stock pricing function, $f(y,t)$, given in proposition 3.3.3. Nevertheless, as explained in section 3.3, this inverse function may be computed without issue using the Newton-Raphson algorithm, and can be cached in order to ensure rapid evaluation.

Once the calibration is complete, the first set of mixture models (based on the un-shifted BSW process) is used to examine the accuracy of our approximate formula, given in proposition 4.2.1, for the local volatility function. Looking at table 4.1, we see that the approximation performs quite well when $\rho = 0$, in that the Monte Carlo call prices implied by the approximate local volatility function are close to the exact analytical prices implied by the mixture model. However, the approximation breaks down for $\rho = 0.4$ and $\rho = -0.4$. One solution to this problem is to re-calibrate the models to market call prices that have been adjusted to account for the observed difference between the Monte Carlo prices and the mixture model prices. In other words, if $C_{mkt}, C_{mc}$ and $C_{mix}$ are the market prices, Monte Carlo prices and mixture model prices respectively, then we calibrate to the adjusted prices

$$C_{adj} = \frac{C_{mkt}C_{mix}}{C_{mc}}.$$ 

Thus, if the ratio between the Monte Carlo prices and the mixture model prices remains constant, our final Monte Carlo prices will agree with the true market prices. Looking at table 4.2, we see that this procedure significantly reduces the magnitude of the errors for $\rho = -0.4$, with relative pricing errors of less than 0.072%, and absolute implied volatility errors of less than 0.8 basis points. Furthermore, the pricing errors were less than the Monte Carlo standard errors across all strikes.

Conversely, for $\rho = 0.4$ the pricing errors remain greater than the standard errors, with relative pricing errors of up to 0.72%, and absolute implied volatility errors of up to 5.7 basis points. However, we can significantly improve these results by repeating the adjustment procedure a second time. Looking at table 4.3, we see that the relative pricing errors fall to less than 0.12%, while the absolute implied volatility errors fall to less than 1.6 basis points. Note that, using the methodology of section 4.4, a comparison of the Monte Carlo prices and the analytical prices for the shifted exponential mixture models and the modified exponential mixture models was also performed. In all cases the Monte Carlo prices agreed with the mixture model prices, up to the standard error of the simulation.

Now, using our third set of calibrated mixture models (based on modified exponential component models) we examine the prices of at-the-money daily-monitored up-and-out call options. We allow the barrier to range between 10% above the initial index value to 100% above this value, in 10%
### Table 4.1: Monte Carlo prices computed using the approximate local volatility function given in proposition 4.2.1

<table>
<thead>
<tr>
<th>Strike</th>
<th>$\rho = -0.4$</th>
<th>$\rho = 0$</th>
<th>$\rho = 0.4$</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>5600</td>
<td>1456.7 (0.59)</td>
<td>1475.2 (0.60)</td>
<td>1497.1 (0.60)</td>
<td>1474.2</td>
</tr>
<tr>
<td>5800</td>
<td>1352.5 (0.58)</td>
<td>1371.6 (0.58)</td>
<td>1394.0 (0.59)</td>
<td>1370.5</td>
</tr>
<tr>
<td>6000</td>
<td>1252.6 (0.56)</td>
<td>1272.3 (0.56)</td>
<td>1295.1 (0.57)</td>
<td>1271.1</td>
</tr>
<tr>
<td>6400</td>
<td>1066.4 (0.53)</td>
<td>1086.9 (0.53)</td>
<td>1110.1 (0.53)</td>
<td>1085.7</td>
</tr>
<tr>
<td>6600</td>
<td>980.47 (0.51)</td>
<td>1001.1 (0.51)</td>
<td>1024.4 (0.52)</td>
<td>999.87</td>
</tr>
<tr>
<td>6700</td>
<td>939.30 (0.50)</td>
<td>960.01 (0.50)</td>
<td>983.21 (0.51)</td>
<td>958.73</td>
</tr>
<tr>
<td>6800</td>
<td>899.33 (0.49)</td>
<td>920.08 (0.49)</td>
<td>943.19 (0.50)</td>
<td>918.78</td>
</tr>
<tr>
<td>6900</td>
<td>860.59 (0.48)</td>
<td>881.34 (0.48)</td>
<td>904.33 (0.49)</td>
<td>880.01</td>
</tr>
<tr>
<td>7000</td>
<td>823.05 (0.47)</td>
<td>843.77 (0.48)</td>
<td>866.62 (0.48)</td>
<td>842.42</td>
</tr>
<tr>
<td>8000</td>
<td>511.67 (0.38)</td>
<td>530.37 (0.39)</td>
<td>550.63 (0.40)</td>
<td>528.99</td>
</tr>
<tr>
<td>9400</td>
<td>240.81 (0.27)</td>
<td>253.71 (0.28)</td>
<td>268.01 (0.28)</td>
<td>252.50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strike</th>
<th>$\rho = -0.4$</th>
<th>$\rho = 0$</th>
<th>$\rho = 0.4$</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>5600</td>
<td>-1.18% (-17.4)</td>
<td>0.068% (1.00)</td>
<td>1.56% (22.9)</td>
<td>-</td>
</tr>
<tr>
<td>5800</td>
<td>-1.32% (-18.0)</td>
<td>0.077% (1.06)</td>
<td>1.71% (23.5)</td>
<td>-</td>
</tr>
<tr>
<td>6000</td>
<td>-1.46% (-18.5)</td>
<td>0.087% (1.10)</td>
<td>1.88% (23.9)</td>
<td>-</td>
</tr>
<tr>
<td>6400</td>
<td>-1.77% (-19.2)</td>
<td>0.110% (1.20)</td>
<td>2.25% (24.5)</td>
<td>-</td>
</tr>
<tr>
<td>6600</td>
<td>-1.94% (-19.4)</td>
<td>0.125% (1.24)</td>
<td>2.45% (24.5)</td>
<td>-</td>
</tr>
<tr>
<td>6700</td>
<td>-2.03% (-19.4)</td>
<td>0.133% (1.28)</td>
<td>2.55% (24.5)</td>
<td>-</td>
</tr>
<tr>
<td>6800</td>
<td>-2.12% (-19.4)</td>
<td>0.142% (1.31)</td>
<td>2.66% (24.4)</td>
<td>-</td>
</tr>
<tr>
<td>6900</td>
<td>-2.21% (-19.4)</td>
<td>0.151% (1.33)</td>
<td>2.76% (24.3)</td>
<td>-</td>
</tr>
<tr>
<td>7000</td>
<td>-2.30% (-19.4)</td>
<td>0.161% (1.35)</td>
<td>2.87% (24.2)</td>
<td>-</td>
</tr>
<tr>
<td>8000</td>
<td>-3.27% (-17.3)</td>
<td>0.261% (1.38)</td>
<td>4.09% (21.6)</td>
<td>-</td>
</tr>
<tr>
<td>9400</td>
<td>-4.63% (-11.7)</td>
<td>0.478% (1.21)</td>
<td>6.14% (15.5)</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strike</th>
<th>$\rho = -0.4$</th>
<th>$\rho = 0$</th>
<th>$\rho = 0.4$</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>5600</td>
<td>21.31 (-0.37)</td>
<td>21.69 (0.001)</td>
<td>22.13 (0.45)</td>
<td>21.68</td>
</tr>
<tr>
<td>5800</td>
<td>21.00 (-0.37)</td>
<td>21.38 (0.003)</td>
<td>21.82 (0.44)</td>
<td>21.37</td>
</tr>
<tr>
<td>6000</td>
<td>20.69 (-0.37)</td>
<td>21.07 (0.005)</td>
<td>21.51 (0.44)</td>
<td>21.06</td>
</tr>
<tr>
<td>6400</td>
<td>20.10 (-0.37)</td>
<td>20.47 (0.008)</td>
<td>20.90 (0.44)</td>
<td>20.47</td>
</tr>
<tr>
<td>6600</td>
<td>19.82 (-0.36)</td>
<td>20.19 (0.010)</td>
<td>20.61 (0.43)</td>
<td>20.18</td>
</tr>
<tr>
<td>6700</td>
<td>19.68 (-0.36)</td>
<td>20.05 (0.011)</td>
<td>20.47 (0.43)</td>
<td>20.04</td>
</tr>
<tr>
<td>6800</td>
<td>19.55 (-0.36)</td>
<td>19.92 (0.012)</td>
<td>20.34 (0.43)</td>
<td>19.91</td>
</tr>
<tr>
<td>6900</td>
<td>19.42 (-0.36)</td>
<td>19.79 (0.013)</td>
<td>20.20 (0.43)</td>
<td>19.78</td>
</tr>
<tr>
<td>7000</td>
<td>19.29 (-0.36)</td>
<td>19.66 (0.014)</td>
<td>20.07 (0.42)</td>
<td>19.65</td>
</tr>
<tr>
<td>8000</td>
<td>18.17 (-0.33)</td>
<td>18.53 (0.020)</td>
<td>18.91 (0.40)</td>
<td>18.51</td>
</tr>
<tr>
<td>9400</td>
<td>16.97 (-0.29)</td>
<td>17.28 (0.023)</td>
<td>17.62 (0.36)</td>
<td>17.26</td>
</tr>
</tbody>
</table>

1. The top third shows MC prices, with standard errors in brackets.
2. The middle third shows the percentage differences between the MC prices and model prices, with the absolute differences in brackets.
3. The bottom third shows the implied volatilities, with the absolute differences between the MC values and model values in brackets.
4. All prices were computed using the same set of random numbers.
Table 4.2: Monte Carlo prices, computed using the approximate local volatility function given in proposition 4.2.1, after fitting to adjusted market prices

<table>
<thead>
<tr>
<th>Strike</th>
<th>$\rho = -0.4$</th>
<th>$\rho = 0$</th>
<th>$\rho = 0.4$</th>
<th>Market</th>
</tr>
</thead>
<tbody>
<tr>
<td>5600</td>
<td>1474.0 (0.60)</td>
<td>1474.2 (0.60)</td>
<td>1476.3 (0.60)</td>
<td>1474.0</td>
</tr>
<tr>
<td>5800</td>
<td>1370.4 (0.58)</td>
<td>1370.6 (0.58)</td>
<td>1372.8 (0.58)</td>
<td>1370.5</td>
</tr>
<tr>
<td>6000</td>
<td>1271.0 (0.57)</td>
<td>1271.2 (0.56)</td>
<td>1273.5 (0.56)</td>
<td>1271.0</td>
</tr>
<tr>
<td>6400</td>
<td>1085.6 (0.53)</td>
<td>1085.7 (0.53)</td>
<td>1088.3 (0.53)</td>
<td>1085.5</td>
</tr>
<tr>
<td>6600</td>
<td>999.76 (0.51)</td>
<td>999.92 (0.51)</td>
<td>1002.5 (0.51)</td>
<td>1000.0</td>
</tr>
<tr>
<td>6700</td>
<td>958.64 (0.51)</td>
<td>958.79 (0.50)</td>
<td>961.38 (0.50)</td>
<td>959.00</td>
</tr>
<tr>
<td>6800</td>
<td>918.69 (0.50)</td>
<td>918.85 (0.49)</td>
<td>921.43 (0.49)</td>
<td>919.00</td>
</tr>
<tr>
<td>6900</td>
<td>879.93 (0.49)</td>
<td>880.08 (0.48)</td>
<td>882.66 (0.48)</td>
<td>879.50</td>
</tr>
<tr>
<td>7000</td>
<td>842.34 (0.48)</td>
<td>842.50 (0.47)</td>
<td>845.07 (0.48)</td>
<td>842.50</td>
</tr>
<tr>
<td>8000</td>
<td>528.90 (0.38)</td>
<td>529.08 (0.39)</td>
<td>531.49 (0.39)</td>
<td>529.00</td>
</tr>
<tr>
<td>9400</td>
<td>252.32 (0.28)</td>
<td>252.60 (0.28)</td>
<td>254.32 (0.27)</td>
<td>252.50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strike</th>
<th>$\rho = -0.4$</th>
<th>$\rho = 0$</th>
<th>$\rho = 0.4$</th>
<th>Market</th>
</tr>
</thead>
<tbody>
<tr>
<td>5600</td>
<td>-0.002% (-0.03)</td>
<td>0.013% (0.19)</td>
<td>0.157% (2.32)</td>
<td>-</td>
</tr>
<tr>
<td>5800</td>
<td>-0.009% (-0.12)</td>
<td>0.007% (0.10)</td>
<td>0.169% (2.31)</td>
<td>-</td>
</tr>
<tr>
<td>6000</td>
<td>0.000% (0.00)</td>
<td>0.017% (0.21)</td>
<td>0.200% (2.54)</td>
<td>-</td>
</tr>
<tr>
<td>6400</td>
<td>0.006% (0.06)</td>
<td>0.020% (0.22)</td>
<td>0.254% (2.76)</td>
<td>-</td>
</tr>
<tr>
<td>6600</td>
<td>-0.024% (-0.24)</td>
<td>-0.008% (-0.08)</td>
<td>0.250% (2.50)</td>
<td>-</td>
</tr>
<tr>
<td>6700</td>
<td>-0.038% (-0.36)</td>
<td>-0.022% (-0.21)</td>
<td>0.248% (2.38)</td>
<td>-</td>
</tr>
<tr>
<td>6800</td>
<td>-0.034% (-0.31)</td>
<td>-0.017% (-0.15)</td>
<td>0.264% (2.43)</td>
<td>-</td>
</tr>
<tr>
<td>6900</td>
<td>0.049% (0.43)</td>
<td>0.066% (0.58)</td>
<td>0.359% (3.16)</td>
<td>-</td>
</tr>
<tr>
<td>7000</td>
<td>-0.019% (-0.16)</td>
<td>0.000% (0.00)</td>
<td>0.305% (2.57)</td>
<td>-</td>
</tr>
<tr>
<td>8000</td>
<td>-0.019% (-0.10)</td>
<td>0.015% (0.08)</td>
<td>0.470% (2.49)</td>
<td>-</td>
</tr>
<tr>
<td>9400</td>
<td>-0.072% (-0.18)</td>
<td>0.041% (0.10)</td>
<td>0.721% (1.82)</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strike</th>
<th>$\rho = -0.4$</th>
<th>$\rho = 0$</th>
<th>$\rho = 0.4$</th>
<th>Market</th>
</tr>
</thead>
<tbody>
<tr>
<td>5600</td>
<td>21.68 (-0.001)</td>
<td>21.69 (0.004)</td>
<td>21.73 (0.047)</td>
<td>21.68</td>
</tr>
<tr>
<td>5800</td>
<td>21.37 (-0.002)</td>
<td>21.37 (0.002)</td>
<td>21.42 (0.046)</td>
<td>21.37</td>
</tr>
<tr>
<td>6000</td>
<td>21.06 (0.000)</td>
<td>21.07 (0.004)</td>
<td>21.11 (0.049)</td>
<td>21.06</td>
</tr>
<tr>
<td>6400</td>
<td>20.46 (0.001)</td>
<td>20.47 (0.004)</td>
<td>20.51 (0.051)</td>
<td>20.46</td>
</tr>
<tr>
<td>6600</td>
<td>20.18 (-0.004)</td>
<td>20.18 (-0.002)</td>
<td>20.23 (0.046)</td>
<td>20.18</td>
</tr>
<tr>
<td>6700</td>
<td>20.04 (-0.006)</td>
<td>20.04 (-0.004)</td>
<td>20.09 (0.043)</td>
<td>20.05</td>
</tr>
<tr>
<td>6800</td>
<td>19.91 (-0.006)</td>
<td>19.91 (-0.003)</td>
<td>19.96 (0.044)</td>
<td>19.91</td>
</tr>
<tr>
<td>6900</td>
<td>19.77 (0.008)</td>
<td>19.78 (0.010)</td>
<td>19.82 (0.057)</td>
<td>19.77</td>
</tr>
<tr>
<td>7000</td>
<td>19.64 (-0.003)</td>
<td>19.65 (0.000)</td>
<td>19.69 (0.047)</td>
<td>19.65</td>
</tr>
<tr>
<td>8000</td>
<td>18.50 (-0.002)</td>
<td>18.51 (0.002)</td>
<td>18.55 (0.047)</td>
<td>18.51</td>
</tr>
<tr>
<td>9400</td>
<td>17.25 (-0.005)</td>
<td>17.26 (0.002)</td>
<td>17.30 (0.044)</td>
<td>17.26</td>
</tr>
</tbody>
</table>

1. The top third shows MC prices, with standard errors in brackets.
2. The middle third shows the percentage differences between the MC prices and market prices, with the absolute differences in brackets.
3. The bottom third shows the implied volatilities, with the absolute differences between the MC values and market values in brackets.
4. All prices were computed using the same set of random numbers.
4.6 Conclusion

We have presented two different methods for determining a local volatility function from market vanilla option prices under stochastic interest rates. By utilizing a mixture-based approach, they have both the flexibility to fit a wide range of market smiles and the tractability to enable rapid calibration to a given set of prices. The calibration procedure is fast because the prices of options are given by a one-dimensional integral, which is a result of certain restrictions placed on the stock pricing function and interest-rate process. For more general models a higher dimensional integral would have to be numerically evaluated.

The first method provides a straightforward analytical approximation for the local volatility
Table 4.4: Prices of at-the-money up-and-out call options valued at 28 April 2015 and expiring on 18 December 2020, computed using the methodology of section 4.4

<table>
<thead>
<tr>
<th>Barrier</th>
<th>$\rho = -0.4$</th>
<th>$\rho = 0.0$</th>
<th>$\rho = 0.4$</th>
<th>Deterministic</th>
</tr>
</thead>
<tbody>
<tr>
<td>7733.58</td>
<td>9.3785 (0.072)</td>
<td>9.6608 (0.077)</td>
<td>9.9951 (0.078)</td>
<td>9.7520 (0.079)</td>
</tr>
<tr>
<td>8436.64</td>
<td>49.592 (0.213)</td>
<td>51.266 (0.223)</td>
<td>53.344 (0.237)</td>
<td>51.958 (0.211)</td>
</tr>
<tr>
<td>9139.69</td>
<td>133.13 (0.381)</td>
<td>138.56 (0.398)</td>
<td>144.02 (0.426)</td>
<td>139.73 (0.383)</td>
</tr>
<tr>
<td>9842.74</td>
<td>253.90 (0.517)</td>
<td>265.05 (0.544)</td>
<td>275.45 (0.574)</td>
<td>265.43 (0.545)</td>
</tr>
<tr>
<td>10545.8</td>
<td>395.29 (0.676)</td>
<td>412.81 (0.732)</td>
<td>429.25 (0.776)</td>
<td>411.70 (0.721)</td>
</tr>
<tr>
<td>11248.8</td>
<td>541.40 (0.925)</td>
<td>562.68 (0.959)</td>
<td>583.24 (0.985)</td>
<td>560.38 (0.940)</td>
</tr>
<tr>
<td>11951.9</td>
<td>676.99 (1.066)</td>
<td>698.18 (1.049)</td>
<td>720.34 (1.085)</td>
<td>695.97 (1.121)</td>
</tr>
<tr>
<td>12655.0</td>
<td>795.50 (1.292)</td>
<td>813.92 (1.321)</td>
<td>835.16 (1.278)</td>
<td>812.92 (1.258)</td>
</tr>
<tr>
<td>13358.0</td>
<td>891.34 (1.432)</td>
<td>904.90 (1.479)</td>
<td>922.72 (1.486)</td>
<td>906.48 (1.540)</td>
</tr>
<tr>
<td>14061.1</td>
<td>966.06 (1.597)</td>
<td>974.82 (1.606)</td>
<td>988.09 (1.628)</td>
<td>978.39 (1.618)</td>
</tr>
</tbody>
</table>

1 The top half shows the MC prices, with the standard errors in brackets.
2 The bottom half shows the percentage differences between the prices based on stochastic rates and on deterministic rates, with the absolute differences in brackets.
3 All prices were computed using the same set of random numbers.
function, which is the weighted average of the local volatilities of each component diffusion, plus a term to allow for differing drifts. The second method allows volatility to be a function of all the state variables, rather than just the stock price and time, and enables exact calibration of mixture models under correlated stochastic interest rates.

Finally, we have given an empirical example that shows how the approximate local volatility function provided in section 4.2 performs in a real world situation. Furthermore, using a mixture of modified exponential models, we have seen how the price of long-dated path dependent derivatives, specifically up-and-out call options, may be significantly affected by interest rate stochasticity.
Chapter 5

Local and Stochastic Volatility under Stochastic Interest Rates Using Mixture Models

5.1 Introduction

Although the LVSR models constructed in the previous chapter are able to accurately fit the implied volatility surface, they may still be criticized for producing unrealistic volatility dynamics. As observed by Coqueret and Tavin (2016), this means that they may significantly misprice exotic derivatives whose value depends on the random nature of volatility itself, such as forward start options. However, existing approaches to combining both local and stochastic volatility under stochastic interest rates are limited by the lack of computationally tractable calibration algorithms.

In the case of deterministic interest rates, Ramponi (2011) solves this problem by constructing a mixture of stochastic volatility (and regime switching) models. By using the multivariate version of the Fokker-Plank equation, he was able to determine the volatility function consistent with a mixture of component multivariate diffusions, such as the Heston (1993) stochastic volatility model. Thus, it is possible to combine the properties of both a local volatility model and a stochastic volatility model, in the sense that the final volatility function will depend on the stock price and time, as well as the volatility process, and that the entire surface of market call prices can be accurately reproduced.

However, when pricing long-term path-dependent options or hybrid options it is no longer safe to assume that interest rates are deterministic. Unfortunately, under stochastic interest rates we
cannot simply price vanilla options using the risk-neutral density of the stock price because the
numeraire (the bank account) is also stochastic. In chapter 4, we overcame this problem by work-
ing in the \( t \)-forward measure for every point in time, \( t \). Instead of using the Fokker-Plank equa-
tion, which only applies in a single fixed measure, we used the PDE governing the evolution of the
\( t \)-forward joint density at time \( t \), as presented by Overhaus et al. (2007). The resulting volatility
function then depends on the joint density of each component model in the \( \text{variable} \ t \)-forward
measure, rather than the \( \text{fixed} \) risk-neutral measure. Nevertheless, the models proposed in the last
chapter all have the drawback that volatility depends purely on the stock price, short-rate and time,
with no allowance for an additional stochastic volatility factor.

In this chapter we extend our mixture model approach to allow for stochastic volatility. Our pro-
cedure superimposes local volatility onto a Stochastic Volatility Stochastic Rates (SVSR) model by
determining the drift and diffusion coefficients consistent with a given mixture of component SVSR
models in the forward measure. Of course, each of these component models contains the same
three state variables as the main model, namely the stock price, the short-rate, and the volatility.
The resulting Local Stochastic Volatility Stochastic Rates (LSVSR) model consists of three correlated
stochastic processes whose diffusion coefficients are fully state dependent. Note that, unlike the
drift of the stock price process, the drift of the volatility process is not determined by no arbitrage
conditions, and thus may differ between the component models and the main model.

Our expressions for the drift and diffusion coefficients in the main model involve the joint den-
sity of the stock price, short-rate and volatility process for each component model. However, SVSR
models typically do not have density functions with known closed-form expressions. Instead, we
must rely on the three-dimensional Fourier transforms of their characteristic functions. Fortu-
nately, a multidimensional extension of the fractional FFT (fast Fourier transform) algorithm found
in Bailey and Swarztrauber (1991) lets us efficiently compute this transform. Unlike a standard FFT,
the fractional FFT allows for independent control of the input and output grid spacing, meaning
that none of the calculations are wasted on points in the extreme tails of the input characteristic
function or output density. This algorithm is also useful for implementing mixtures of stochastic
volatility models under deterministic interest rates. For example, under the standard Heston
model, Ramponi (2011) indicates that the joint density of the stock price and volatility can be de-
termined from its characteristic function. However, he gives no specific algorithm for doing so, and
does not present any concrete simulation procedures or Monte Carlo tests regarding his proposed
mixture of Heston models.

To provide a numerical example, we will use components based on the Schöbel-Zhu-Hull-White
(SZHW) model presented by Grzelak, Oosterlee, and van Weeren (2012), in which the short-rate
and volatility each follow correlated Ornstein-Uhlenbeck processes. We have extended their results
to find the joint characteristic function under piecewise-constant time-dependent parameters, as
this is useful for fitting multiple expiries. Via the FFT based approach of Carr and Madan (1999),
as updated by Chourdakis (2004) to use the fractional FFT, we calibrate our mixture model to FTSE
100 index option data. Then, employing the multidimensional fractional FFT to efficiently cache
5.1. Introduction

the necessary joint densities, we simulate the resulting process using our expressions for the drift and volatility functions implied by a mixture of SZHW models. This simulation is used to test the accuracy of our procedure, and also study the impact of non-deterministic interest rates on long-term path-dependent options, specifically up-and-out call options.

Naturally, it is possible to construct mixtures using any other SVSR model with a known joint characteristic function, such as the Heston-Hull-White model of Grzelak and Oosterlee (2011), or the Heston-CIR model of Recchioni and Sun (2016). Nonetheless, we have chosen to focus on the SZHW model as it allows for an arbitrary matrix of correlations between the three driving Brownian motions, unlike the aforementioned models.

A number of other authors have looked at stochastic volatility, local volatility and stochastic interest rates. However, they have typically combined at most two out of these three extensions to the standard Black-Scholes model. For example, we refer the reader to Benhamou et al. (2012), Benhamou et al. (2008), Grzelak et al. (2008), Ren et al. (2007) and van der Stoep et al. (2016) for local volatility under stochastic rates. Likewise, for recent work on local and stochastic volatility, we recommend An and Li (2015), Henry-Labordere (2009), Lipton et al. (2014), Lorig et al. (2015), Piterbarg (2007), Ren et al. (2007), Tian, Zhu, Lee, Klebaner, and Hamza (2015) and van der Stoep et al. (2016). Finally, we suggest Grzelak and Oosterlee (2011), Grzelak et al. (2012) and Recchioni and Sun (2016) for some examples of stochastic volatility models under stochastic rates.

One of the very few papers to consider a full LSVSR model is Deelstra and Rayée (2013). However, they only give an explicit formula for the local volatility coefficient in the case of a stochastic volatility process that is independent of the stock price. In the more realistic case of dependent stochastic volatility, they suggest that the model can be calibrated using either Monte Carlo simulation or a numerical PDE based approach. Nevertheless, no actual calibrations or numerical examples are given for either dependent or independent stochastic volatility. Conversely, our mixture-based approach avoids the need to perform any Monte Carlo simulations or to solve any PDEs numerically, which may be quite costly procedures. We also provide an example calibration to real world data, along with all the necessary implementational details, and examine the effect of interest rate stochasticity on barrier option prices.

The remainder of this chapter is organized as follows. In section 5.2 below introduce our proposed mixture of SVSR models, and explain how to use the multidimensional fractional FFT to compute the necessary joint densities. An example calibration to FTSE 100 index option prices, along with an examination of the effect of interest rate stochasticity on barrier option prices, is provided in section 5.3. Finally, we conclude in section 5.4.
5.2 Mixture of Stochastic Volatility Models under Stochastic Interest Rates

In chapter 4 we introduced a new methodology for constructing multivariate mixture models in the variable \( t \)-forward measure. Our main result was a formula for the drift and volatility of a risk-neutral diffusion whose \( t \)-forward joint density equals the weighted average of a given set of component joint densities. Each of these component joint densities were produced by a component diffusion with the same set of state variables as the main diffusion, e.g. the stock price, the short-rate, and the bank account. Under Hull-White interest rates, we applied this result to construct mixtures of parametric local volatility models with analytically known joint densities, specifically the shifted exponential model, the hyperbolic sine model, and the modified exponential model.

The key limitation of this work was that it did not allow for stochastic volatility. As explained by Rebonato (2004), local volatility on its own is not sufficient to accurately price certain path-dependent derivatives, and does not produce realistic smile dynamics or forward volatility curves. Therefore, our objective is to extend this procedure to the case of stochastic volatility by deriving the joint characteristic function of a suitable SVSR component model, and then utilizing the multidimensional fractional FFT to efficiently cache the resulting joint density.

Our proposed model is based on the SZHW process described by Grzelak et al. (2012), which we have extended to allow for piecewise-constant time-dependent parameters, as this allows us to accurately calibrate to multiple expiries. We have also included an extra time-dependent volatility multiplier, \( m_{k,t} \), because it adds more flexibility to the mixture-diffusion that will be constructed in theorem 5.2.3. The full details of our SZHW component processes are given below. Note that these individual processes do not include any local volatility term. This term only appears in the associated mixture-diffusion, given in equation (5.2.7).

**Assumption 5.2.1.** Under assumption 4.3.2, let \( X^{(k)}_t := (S_{k,t}, r_{k,t}, \eta_{k,t}) \), be the stock price, short-rate and volatility in component model \( k = 1, 2, ..., n \), such that

\[
\begin{align*}
    dS_{k,t} &= (r_{k,t} - y_t) S_{k,t} dt + m_{k,t} \eta_{k,t} S_{k,t} d\tilde{W}_{1,t}, \\
    dr_{k,t} &= (\theta_t - a_t r_{k,t}) dt + \psi_t d\tilde{W}_{2,t}, \\
    d\eta_{k,t} &= \kappa_{k,t} (\bar{\eta}_{k,t} - \eta_{k,t}) dt + \gamma_{k,t} d\tilde{W}_{3,t},
\end{align*}
\]

where \( y_t, m_{k,t}, \theta_t, a_t, \psi_t, \kappa_{k,t}, \bar{\eta}_{k,t} \) and \( \gamma_{k,t} \) are deterministic functions of \( t \), and \( \tilde{W}_t := (\tilde{W}_{1,t}, \tilde{W}_{2,t}, \tilde{W}_{3,t}) \) is a joint Brownian motion with time-dependent correlation matrix

\[
\rho_{k,t} := \begin{pmatrix}
1 & \rho_{k,t,1,2} & \rho_{k,t,1,3} \\
\rho_{k,t,1,2} & 1 & \rho_{k,t,2,3} \\
\rho_{k,t,1,3} & \rho_{k,t,2,3} & 1
\end{pmatrix},
\]

Also assume that the parameters \( m_{k,t}, a_t, \psi_t, \kappa_{k,t}, \bar{\eta}_{k,t}, \gamma_{k,t}, \rho_{k,t,1,2}, \rho_{k,t,1,3}, \rho_{k,t,2,3} \) are piecewise-
5.2. Mixtures of Stochastic Volatility Models under Stochastic Interest Rates

Looking at the equations above, we see that the short-rate process has mean reversion rate $a_t$, and volatility $\psi_t$. Likewise, the volatility process has mean reversion rate $\kappa_{k,t}$, and volatility $\gamma_{k,t}$. The parameters $\theta_t$ and $\tilde{\eta}_{k,t}$ determine the level to which $r_{k,t}$ and $\eta_{k,t}$ revert, with $\theta_t$ being chosen so that the short-rate model exactly reproduces the current market yield curve. Unlike the other parameters, the dividend rate $y_t$, and the parameter $\theta_t$, do not have to be piece-wise constant.

Note that the instantaneous covariance matrix of $X_{t}^{(k)}$ depends on $\eta_{k,t}^2$, and is therefore not an affine function of the state variables. Thus, in order to derive the necessary characteristic function using the results of Duffie et al. (2000), we include $v_{k,t} := \eta_{k,t}^2$ as an additional state variable. We will also simplify the derivation by transforming $S_{k,t}$ and $r_{k,t}$ as described below.

**Definition 5.2.1.** Let $Y_{t}^{(k)} := (\tilde{x}_{k,t}, \tilde{r}_{k,t}, v_{k,t}, \eta_{k,t})$, where $\tilde{x}_{k,t} := \log(S_{k,t}) + D_t - \Upsilon_t$, $\tilde{r}_{k,t} := r_{k,t} - \xi_t$, $v_{k,t} := \eta_{k,t}^2$, and

\[
\xi_t := f_{mk}(0, t) + \int_0^t h(s, t)H(s, t)ds, \quad \Upsilon_t := -\log(P(0, t)) + \frac{1}{2} \int_0^t H^2(u, t)du,
\]

\[
D_t := \int_0^t y_u du, \quad h(u, v) := \psi_u e^{-\frac{v}{a_u}dz}, \quad H(u, t) := \int_u^t h(u, v)dv.
\]

Here $\xi_t$ is the expected value of the short-rate in the risk-neutral measure under the Hull-White model, and $\Upsilon_t$ is the integral of $\xi_u$ over $u$ from 0 to $t$. We refer the reader to the appendix of chapter 3 for a derivation of these two functions, and for further details regarding the Hull-White model.

Existing literature on the SZHW model (and most other hybrid stochastic volatility models) only gives the marginal characteristic function of the log stock price, as this is sufficient to price vanilla options. Conversely, we derive the joint characteristic function of the log stock price, short-rate and volatility processes, since this is necessary to compute the joint density function that appears in theorem 4.3.2. See appendix 5.A for the proof.

**Theorem 5.2.1.** Under assumption 5.2.1, let $T$ be a fixed expiry time, and define $\tau := T - t$. Then, for $\tau \in (\tau_j, \tau_{j+1}]$, we can write the piecewise-constant parameters of each component model as $m_{k,T-\tau} = \tilde{m}_{k,j}$, $a_{T-\tau} = \tilde{a}_j$, and so on, where $0 = \tau_0 < \tau_1 < \ldots < \tau_p = T$. Next, recursively define the following functions for $u := (u_1, u_2, u_3, u_4) \in \mathbb{C}^4$ and $\tau \in (\tau_j, \tau_{j+1}]$:

\[
B_k(u, \tau) = (B_{k,1}(u, \tau), B_{k,2}(u, \tau), B_{k,3}(u, \tau), B_{k,4}(u, \tau)),
\]

\[
B_{k,1}(u, \tau) = iu_1, \quad B_{k,2}(u, \tau) = b + (B_{k,2,j} - b) e^{-\tilde{a}_j(\tau - \tau_j)},
\]

\[
B_{k,3}(u, \tau) = B_{k,3,j} + \left( \frac{1 - e^{d(\tau - \tau_j)}}{1 - ge^{d(\tau - \tau_j)}} \right) h, \quad B_{k,4}(u, \tau) = \left( \frac{1 - g}{1 - ge^{d(\tau - \tau_j)}} \right) (B_{k,4,j} + f_5),
\]

\[
A_k(u, \tau) = A_{k,j} + \tilde{\gamma}^2_{k,j} \left( (B_{k,3,j} + h)(\tau - \tau_j) + \frac{(1 - g)h}{gd} \log \left( \frac{1 - ge^{d(\tau - \tau_j)}}{1 - g} \right) \right)
\]
Local and Stochastic Volatility under Stochastic Interest Rates Using Mixture Models

Then the

\[ f_5 = 2\tilde{\kappa}_{k,j}\tilde{\eta}_{k,j}(B_{k,3,j}f_1 + h f_2) + \tilde{\psi}_{j}\tilde{\eta}_{k,j}\tilde{\rho}_{k,j,1,2}i u_1 (b f_1 + (B_{k,2,j} - b) (f_3 + g f_4)) + 2\tilde{\psi}_{j}\tilde{\kappa}_{k,j}\tilde{\rho}_{k,j,2,3}(b (B_{k,3,j}f_1 + h f_2) + (B_{k,2,j} - b) ((B_{k,3,j} + h) f_3 + (gB_{k,3,j} + h) f_4)), \]

\[ f_6 = \int_\tau \left( \tilde{\kappa}_{k,j}\tilde{\eta}_{k,j} + \tilde{\psi}_{j}\tilde{\kappa}_{k,j}\tilde{\rho}_{k,j,2,3}B_{k,2}(u, y) + \frac{1}{2}\tilde{\kappa}_{k,j}^2B_{k,4}(u, y) \right) B_{k,4}(u, y) \, dy. \]

Then the \( T \)-forward characteristic function of \( Y_T^{(k)} \) is

\[
\hat{\phi}_{Y_T^{(k)}}(u) = \mathbb{E}^T\left(e^{iuY_T^{(k)}}\right) = P(0, T)^{-1}e^{-Y_T + A_k(u, T) + B_k(u, T)Y_0^{(k)}}.
\]

Finally, let \( \tilde{X}_T^{(k)} := (\tilde{x}_{k,t}, \tilde{r}_{k,t}, \eta_{k,t}) \) and \( w := (w_1, w_2, w_3) \in \mathbb{C}^3. \) Then the joint PDF of \( X_T^{(k)} \) in the \( T \)-forward measure is

\[
\phi_{k,T}(x) = \frac{1}{x_1}\phi_{Y_T^{(k)}}(\log(x_1) + D_T - Y_T, t_2 - \xi_T, t_3),
\]

where the joint PDF of \( \tilde{X}_T^{(k)} \) is equal to the inverse Fourier transform of the function \( \hat{\phi}_{X_T^{(k)}}(w) = \hat{\phi}_{Y_T^{(k)}}(w_1, w_2, 0, w_3), \)

\[
\phi_{\tilde{X}_T^{(k)}}(x) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} e^{-iw \cdot x} \hat{\phi}_{X_T^{(k)}}(w) \, dw. \quad (5.2.2)
\]

Note that evaluating the definite integral \( f_6 \) requires computation of the ordinary hypergeometric function, \( _2F_1 \), which does not have a closed-form solution. The analysis performed by Grzelak et al. (2012) suggests that \( f_6 \) be evaluated using Simpson's rule with an evenly spaced partition of
size 64.

We are also required to numerically invert the characteristic function of $Y^{(k)}_T$ with respect to the first, second and fourth inputs. This can be computed efficiently on a three-dimensional grid of points, evenly spaced in each dimension, using an FFT. However, some care needs to be taken because a standard FFT imposes an inverse relationship between the spacing in the input and output grids. This means that using a fine grid to integrate the characteristic function, as is necessary to achieve a high degree of accuracy, results in a density function defined on a grid that is far too sparse.

Fortunately, this problem can be overcome using the fractional FFT, as described by Bailey and Swarztrauber (1991). The idea is that instead of using a standard FFT to compute sums of the form

$$D_k(x) := \sum_{j=0}^{m-1} e^{-2\pi ijk/m} x_j, \quad k = 0, 1, 2, ..., m - 1,$$

we can use a fractional FFT to compute sums of the form

$$G_k(x, \alpha) := \sum_{j=0}^{m-1} e^{-2\pi ijk\alpha} x_j, \quad k = 0, 1, 2, ..., m - 1,$$

for a given complex number $\alpha$. An extension of the fractional FFT algorithm in Bailey and Swarztrauber (1991) to multiple dimensions yields the following theorem.

**Theorem 5.2.2.** Let $Z_n := \{0, 1, ..., n-1\}$, $n := (n_1, n_2, ..., n_d) \in \mathbb{N}^d$, $[n] := \prod_{j=1}^{d} n_j$, $Z_n := Z_{n_1} \times Z_{n_2} \times ... \times Z_{n_d}$, and $\mathbb{C}^n := \mathbb{C}^{n_1 \times n_2 \times \cdot \cdot \cdot \times n_d}$. Furthermore, let $D(X) \in \mathbb{C}^n$ be the discrete Fourier transform (DFT) of matrix $X \in \mathbb{C}^n$, and $D^{-1}(X) \in \mathbb{C}^n$ be the inverse DFT, with elements

$$D_k(X) := \sum_{j=0}^{n-1} e^{-2\pi i j \cdot (k/n)} X_j, \quad D_k^{-1}(X) := \frac{1}{[n]} \sum_{j=0}^{n-1} e^{2\pi i j \cdot (k/n)} X_j, \quad k \in Z_n,$$

where $X_j$ is the $(j_1, j_2, ..., j_d)$ element of $X$, for multi-index $j \in Z_m$, and `/` denotes element-wise division. Now, given $\alpha \in \mathbb{C}^d$, $\mathbf{m} \in \mathbb{N}^d$ and $\mathbf{H} \in \mathbb{C}^m$, define the matrices $\mathbf{Y}$ and $\mathbf{Z} \in \mathbb{C}^{2m}$ such that

$$\mathbf{Y}_j := \begin{cases} e^{-i\pi j \cdot (j+\alpha)} \mathbf{H}_j, & j \in Z_m \\ 0, & j \in Z_{2m} \setminus Z_m \end{cases}, \quad \mathbf{Z}_j := \prod_{n=1}^{d} b_n(j_n), \quad j \in Z_{2m},$$

$$b_n(j) := \begin{cases} e^{i\pi \alpha_n}, & j \in Z_{m_n} \\ e^{i(2m_n - j) \alpha_n}, & j \in Z_{2m_n} \setminus Z_{m_n} \end{cases},$$

where `*` denotes element-wise multiplication. Then the fractional Fourier transform of $\mathbf{H}$, $G(\mathbf{H}, \alpha) \in \mathbb{C}^{2m}$.
$\mathbb{C}^m$, consists of the elements

$$G_k(X, \alpha) := \sum_{j=0}^{m-1} e^{-2\pi i j (k \cdot \alpha)} H_j$$

$$= e^{-i \pi k \cdot \alpha} D_k^{-1} (D(Y) \ast D(Z)), \quad k \in \mathbb{Z}_m. \quad (5.2.3)$$

Note that we discard the terms $D_k^{-1} (D(Y) \ast D(Z))$ for all $k \in \mathbb{Z}_{2m} \setminus \mathbb{Z}_m$. Using equation (5.2.3), a fractional Fourier transform of size $[m]$ can be computed using three FFTs of size $[2m]$. It therefore requires only $O([m] \log([m]))$ operations, and can be implemented using the wide range of highly optimized FFT algorithms currently available.

We now have all the tools necessary to evaluate the RHS of equation (5.2.2) efficiently. Firstly, given a size vector $m \in \mathbb{N}^3$, let us define the input grid $w_j = w_0 + j \cdot g$, where $w_0 \in \mathbb{R}^3$, $j \in \mathbb{Z}_m$ and $g \in \mathbb{R}^3$ are the grid’s starting point, multi-index and step-size respectively. Likewise, define the output grid $x_k = x_0 + k \cdot h$. Then, omitting the subscript $\tilde{X}^{(k)}_T$ for clarity, we have

$$\phi(x_k) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} e^{-i w \cdot x_k} \hat{\phi}(w) \, dw$$

$$\approx \frac{2|g|}{(2\pi)^3} \sum_{j=0}^{m-1} e^{-i w_j \cdot x_k} \hat{\phi}(w_j)$$

$$= \frac{2|g|}{(2\pi)^3} e^{-i w_0 \cdot x_k} \sum_{j=0}^{m-1} e^{-i j \cdot g \cdot k \cdot h} \hat{\phi}(w_j) e^{-i j \cdot g \cdot x_0}, \quad (5.2.5)$$

where (5.2.4) is due to the fact that $\Re(e^{-i w \cdot x_k} \hat{\phi}(w)) = \Re(e^{i w \cdot x_k} \hat{\phi}(-w))$ for characteristic functions of real random variables. Note that the input grid should be chosen to cover a large region of $\mathbb{R}^+ \times \mathbb{R}^2$, depending on the decay of $\hat{\phi}(w)$. Likewise, the output grid should be chosen to cover a large region of $\mathbb{R}^3$, depending on the mean and variance of $\tilde{X}^{(k)}_T$, which can be deduced from $\hat{\phi}(w)$.

The summation in line (5.2.5) is a fractional Fourier transform with $2\pi \alpha = g \ast h$ and $H_j = \hat{\phi}(w_j) e^{-i j \cdot g \cdot x_0}$. Comparing this with the standard Fourier transform, we see that the use of a fractional FFT has allowed us to avoid setting $g \ast h = 2\pi / m$. Thus we are able to independently control the spacing of the input and output grids, which justifies the additional computational cost compared to a standard FFT. Also observe that we can easily accommodate a more accurate quadrature rule, such as Simpson’s rule, by multiplying $H_j$ by the appropriate factors.

Our next task is to calibrate the mixture model, one expiry at a time, by minimizing the sum of squared differences between the market prices of vanilla options and their weighted average model.
5.2. Mixtures of Stochastic Volatility Models under Stochastic Interest Rates

prices. Given the $T$-forward characteristic function of $x_{k,T} := \log (S_{k,T})$,

$$
\hat{\phi}_{x_{k,T}}(w) = \mathbb{E}^T (e^{iw x_{k,T}}) \\
= \mathbb{E}^T (e^{iw(\hat{x}_{k,T}+Y_T-D_T)}) \\
= e^{iw(Y_T-D_T)} \hat{\phi}_{\hat{X}_T} (w, 0, 0),
$$

(5.2.6)

we can price vanilla options on an evenly spaced set of log-strikes using the FFT based approach of Carr and Madan (1999). Specifically, we evaluate the relevant complex line integral using the contour with constant imaginary part of $1/2$, which is the middle of the domain of validity for a covered call (from which we get calls and puts). We refer the reader to Joshi (2011) for a detailed discussion of this approach, along with some improvements, and to Mrázek, Pospíšil, and Sobotka (2016) for an analysis of suitable optimization algorithms. Also note that, as explained by Chourdakis (2004), the one-dimensional fractional FFT can be used here to ensure that the output set of log-strikes is not too sparse. We now have all the ingredients necessary to construct our main model using theorem 4.3.2.

**Theorem 5.2.3.** Under assumptions 4.3.1 and 5.2.1, let $X_t := (S_t, r_t, \eta_t)$, be the stock price, short-rate and volatility. Also let

$$
\begin{align*}
\dot{S}_t &= (r_t - y_t) S_t dt + \nu (X_t, t) \eta_t S_t dW_{1,t}, \\
\dot{r}_t &= (\theta_t - \alpha_t r_t) dt + \psi_t \left( \rho_{1,2} (X_t, t) dW_{1,t} + \sqrt{1 - \rho_{1,2}^2} (X_t, t) dW_{2,t} \right), \\
\dot{\eta}_t &= \mu_3 (X_t, t) dt + \sigma_{1,3} (X_t, t) dW_{1,t} + \sigma_{2,3} (X_t, t) dW_{2,t} + \sigma_{3,3} (X_t, t) dW_{3,t},
\end{align*}
$$

(5.2.7)

where $\nu (x, t), \rho_{1,2} (x, t), \mu_3 (x, t), \sigma_{1,3} (x, t), \sigma_{2,3} (x, t)$ and $\sigma_{3,3} (x, t)$ are given by

$$
\nu (x, t) = \sqrt{\Lambda \left( m_t * m_t \right)}, \quad \rho_{1,2} (x, t) = \frac{\Lambda \left( m_t * \rho_{1,2} \right)}{\nu (x, t)},
$$

$$
\mu_3 (x, t) = \Lambda \left( \kappa_t * (\bar{\eta}_t - x_3) \right), \quad \sigma_{1,3} (x, t) = \frac{\Lambda \left( m_t * \gamma_t * \rho_{1,1,3} \right)}{\nu (x, t)},
$$

$$
\sigma_{2,3} (x, t) = \frac{\Lambda \left( \gamma_t * \rho_{1,2,3} \right) - \rho_{1,2} (x, t) \sigma_{1,3} (x, t)}{\sqrt{1 - \rho_{1,2}^2} (x, t)},
$$

$$
\sigma_{3,3} (x, t) = \sqrt{\Lambda \left( \gamma_t * \gamma_t \right) - \sigma_{1,3}^2 (x, t) - \sigma_{2,3}^2 (x, t)},
$$

where the $n$-dimensional vectors $m_t, \kappa_t, \bar{\eta}_t, \gamma_t, \rho_{1,1,2}, \rho_{1,1,3}$ and $\rho_{1,2,3}$ contain the parameters of the $n$ component models, and the function $\Lambda \left( c_t \right) = \Lambda \left( c_t, x, t \right)$ is defined as

$$
\Lambda \left( c_t, x, t \right) := \frac{\sum_{k=1}^{n} \lambda_k c_{k,t} \phi_{k,t} (x)}{\sum_{k=1}^{n} \lambda_k \phi_{k,t} (x)}.
$$
Then, under the assumptions of theorem 4.3.2,

\[ \phi_t(x) = \sum_{k=1}^{n} \lambda_k \phi_{k,t}(x). \]

Note that the coefficients \( \nu, \rho_{1,2}, \sigma_{1,3}, \sigma_{2,3}, \sigma_{3,3} \) defined above were determined by taking the Cholesky decomposition of the matrix \( \Sigma \Sigma' \) defined in theorem 4.3.2.

In order to gain some insight into this model, suppose that all the component volatility processes are governed by the same set of parameters, i.e., \( \kappa_{k,t} = \kappa_t, \bar{\eta}_{k,t} = \bar{\eta}_t \) and \( \gamma_{k,t} = \gamma_t \) for all \( k \). Since, by assumption 4.3.2, the initial value of volatility must be the same for each component model, this means that the volatility multipliers, \( m_{k,t} \), and possibly the instantaneous correlations, \( \rho_{k,t,1,3} \), are the only parameters allowed to vary with \( k \). Thus, letting \( (\hat{W}_{1,t}, \hat{W}_{2,t}, \hat{W}_{3,t}) \) be a joint Brownian motion with a state dependent correlation matrix, the model reduces to the form

\[
\begin{align*}
    dS_t &= (r_t - y_t)S_t dt + \nu(X_{t, t}, t) \eta_t S_t d\hat{W}_{1,t}, \\
    dr_t &= (\theta_t - a_t r_t) dt + \psi_t d\hat{W}_{2,t}, \\
    d\eta_t &= \kappa_t (\bar{\eta}_t - \eta_t) dt + \gamma_t d\hat{W}_{3,t},
\end{align*}
\]

where \( \nu^2 \) is a (fully state dependent) weighted average of \( m_{k,t}^2 \). We see that, in this case, the local volatility function, \( \nu \), has been superimposed onto the underlying SVSR model. However, unlike in a traditional LSVSR model, \( \nu \) depends on the short-rate and volatility, in addition to the stock price and time.

The key parameters determining the stochastic nature of volatility are \( m_{k,t}, \kappa_{k,t}, \bar{\eta}_{k,t}, \gamma_{k,t} \) and \( \rho_{k,t,1,3} \). Observe that setting \( m_{k,t} \) close to 1 for all \( k \) reduces the influence of the local volatility function, \( \nu \). This means that the model’s implied volatility smile must be primarily generated by the stochasticity of \( \eta_t \). On the other hand, setting \( \gamma_{k,t} \) close to 0 moves \( \eta_{k,t} \) towards a deterministic function of time. If this is done for all the components, then we are back in the case of local volatility. Thus, by restricting the parameters \( m_{k,t} \) or \( \gamma_{k,t} \) accordingly, it is possible to control the balance between local and stochastic volatility.

Although we focus on the case of the SZHW model throughout this chapter, the overall approach can also be used to calibrate mixtures of other SVSR models as long as the joint characteristic function of the log stock price, short-rate, and volatility is known. In other words, the derivation of theorem 5.2.3 does not depend on the specific choice of interest rate process or volatility process, and equation 5.2.5 can be used to compute the necessary joint density as long as the aforementioned characteristic function is available.

Compared to the mixture models presented in the previous chapter, our model has the key advantage that volatility is not entirely determined by the stock price, short-rate and bank account, and may therefore have more realistic dynamics. It also avoids negative stock prices, allows for significant skew in the Black-Scholes implied volatility surface, and can be calibrated one expiry at
5.3 Empirical Results

In this section we will perform an example calibration to real market quotes as at \( t_0 = 28 \) April 2015. Using the same data as sections 3.4 and 4.5, so that our results are directly comparable, we will see what sort of fits can be obtained, and also examine the effect of interest rate stochasticity on long-term path-dependent options, specifically up-and-out call options.

We begin by fitting the model given in theorem 5.2.3 to FTSE 100 index call options expiring on \( t_1 = 15 \) December 2017 and \( t_2 = 18 \) December 2020. We have chosen to use two equally weighted component models because this was sufficient to provide an almost exact fit to market prices. Given that the instantaneous correlation between the short-rate and the stock price is not implied by vanilla option prices, we perform three separate calibration for \( \rho_{1,2} = -0.4 \), \( \rho_{1,2} = 0 \), and \( \rho_{1,2} = 0.4 \). As explained by Hörfelt and Torné (2010), historical correlation is typically within this range. We also assume that the instantaneous correlation between the short-rate and volatility, \( \rho_{2,3} \), is zero. Note that we have dropped the subscripts \( k \) and \( t \) because we have assumed that these interest rate correlations are the same for all component models, \( k \), and times, \( t \).

The remaining parameters, \( m_{k,t}, \kappa_{k,t}, \bar{\eta}_{k,t}, \gamma_{k,t} \) and \( \rho_{k,t,1,3} \), are assumed to be constant on the intervals \([t_0, t_1)\) and \([t_1, t_2)\). We perform the calibration one expiry at a time, starting with the first. In this way, by working progressively forward in time, it is possible to accommodate any number of expiries. Note that the initial value of volatility, \( \eta_0 \), is not observable in the market, so must be included as an additional parameter for the first expiry.

The quality of the fit is shown in figure 5.1. Note that we have only shown results for \( \rho_{1,2} = 0.4 \) as the results for other values of \( \rho_{1,2} \) were very similar. We see that our mixture model is able to provide a very good fit for both expiries, even with only two components. Unlike the mixture of
shifted exponential models presented in chapter 4, we also see that it is able reproduce the extreme skew present at the December 2017 expiry.

Next, as was done in section 4.5, we examine the prices of at-the-money daily-monitored up-and-out call options. We again allow the barrier to range between 10% above the initial index value to 100% above this value, in 10% increments, so that our results can be directly compared to those shown in table 4.4. Prices are calculated for each value of correlation via Monte Carlo simulation of SDE (5.2.7), using an Euler discretization scheme with daily stepping and 400,000 paths. The joint densities, $\phi_{k,t}$, necessary to compute the coefficients of this SDE are evaluated using a fractional FFT via equation (5.2.5). More specifically, for both component models we cache the value of $\phi_{k,t}$ on an array with dimension $480 \times 480 \times 480$, and for 100 evenly spaced times between time 0 and the final expiry. The array is chosen to cover 5 standard deviations above and below the mean of each state variable. The required value of $\phi_{k,t}$ at each step of the simulation is then computed by linearly interpolating between the neighbouring $2^4$ elements of the cache.

Comparing table 5.1 to table 4.4, we see that the prices of barrier options are significantly affected by the inclusion of stochastic volatility, with higher prices observed across the board in table 5.1. Furthermore, we again see that the instantaneous correlation between the stock price and short-rate can have a large impact on the results. The difference between the negative and positive correlation cases reaches up to 6.93%, with lower prices under negative correlation and higher prices under positive correlation. As noted in section 4.5, barrier option prices under deterministic rates will be very similar to those of the zero correlation case, meaning that the impact of interest rate stochasticity is most important when correlation is non-zero.

In order to assess the accuracy of our new calibration and simulation procedure, we have also provided, in table 5.2, the Monte Carlo prices of vanilla options for the final expiry. As with figure...
5.4. Conclusion

Table 5.2: Prices of out-of-the-money vanilla options for $\rho_{1,2} = 0.4$, valued at 28 April 2015 and expiring on 18 December 2020, computed using Monte Carlo simulation of the model given in theorem 5.2.3

<table>
<thead>
<tr>
<th>Strike</th>
<th>MC Price$^1$</th>
<th>Mkt Price</th>
<th>Price Error$^2$</th>
<th>MC IV</th>
<th>Mkt IV</th>
<th>IV Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>5600</td>
<td>819.17 (1.71)</td>
<td>818.00</td>
<td>0.14% (0.68)</td>
<td>21.71</td>
<td>21.68</td>
<td>0.027</td>
</tr>
<tr>
<td>5800</td>
<td>899.79 (1.81)</td>
<td>899.00</td>
<td>0.09% (0.44)</td>
<td>21.39</td>
<td>21.37</td>
<td>0.018</td>
</tr>
<tr>
<td>6000</td>
<td>984.96 (1.93)</td>
<td>984.00</td>
<td>0.10% (0.49)</td>
<td>21.08</td>
<td>21.06</td>
<td>0.019</td>
</tr>
<tr>
<td>6400</td>
<td>1083.5 (2.65)</td>
<td>1085.5</td>
<td>-0.19% (-0.77)</td>
<td>20.43</td>
<td>20.46</td>
<td>-0.038</td>
</tr>
<tr>
<td>6600</td>
<td>997.97 (2.52)</td>
<td>1000.0</td>
<td>-0.20% (-0.80)</td>
<td>20.15</td>
<td>20.18</td>
<td>-0.037</td>
</tr>
<tr>
<td>6700</td>
<td>956.99 (2.45)</td>
<td>959.00</td>
<td>-0.21% (-0.82)</td>
<td>20.01</td>
<td>20.05</td>
<td>-0.037</td>
</tr>
<tr>
<td>6800</td>
<td>917.16 (2.39)</td>
<td>919.00</td>
<td>-0.20% (-0.77)</td>
<td>19.88</td>
<td>19.91</td>
<td>-0.034</td>
</tr>
<tr>
<td>6900</td>
<td>878.47 (2.32)</td>
<td>879.50</td>
<td>-0.12% (-0.45)</td>
<td>19.75</td>
<td>19.77</td>
<td>-0.019</td>
</tr>
<tr>
<td>7000</td>
<td>840.93 (2.26)</td>
<td>842.50</td>
<td>-0.19% (-0.70)</td>
<td>19.62</td>
<td>19.65</td>
<td>-0.029</td>
</tr>
<tr>
<td>8000</td>
<td>527.16 (1.66)</td>
<td>529.00</td>
<td>-0.35% (-1.11)</td>
<td>18.47</td>
<td>18.51</td>
<td>-0.035</td>
</tr>
<tr>
<td>9400</td>
<td>251.40 (1.13)</td>
<td>252.50</td>
<td>-0.44% (-0.97)</td>
<td>17.23</td>
<td>17.26</td>
<td>-0.027</td>
</tr>
</tbody>
</table>

1 Standard errors are shown in brackets.
2 The absolute difference in terms of standard errors is shown in brackets.
3 Put prices are shown for strike 6000 and below.
4 Call prices are shown for strike 6400 and above.
5 All prices were computed using the same set of random numbers.

5.1, we have only provided results for $\rho_{1,2} = 0.4$ because the results for the other cases were very similar. We see that the Monte Carlo prices do indeed agree with the market prices, with most of the deviations lying within one standard error in terms of price, and 4 basis points in terms of implied volatility.

5.4 Conclusion

In this chapter we have provided an efficient methodology for constructing local and stochastic volatility models under stochastic interest rates. It allows for time-dependent parameters, arbitrary instantaneous correlation between all the state variables, and enables rapid calibration to a wide range of market smiles. When calibrating, we make no approximations, and do not have to rely on numerical PDEs, asymptotic expansions, Markovian projections, or Monte Carlo simulations, which may be slow or of limited applicability.

Our approach involves the inversion of three-dimensional characteristic functions, which can be performed efficiently using the multidimensional fractional FFT. This inversion only needs to be performed once, after the model is calibrated, and not repeatedly during the calibration procedure. Our approach is also quite general, and can be used to construct mixture models based on any set of component models with known characteristic functions.

Lastly, we have given an empirical example based on a mixture of two Schöbel-Zhu-Hull-White components, which demonstrates how well the model performs in a real world situation. This example also shows how correlated stochastic interest rates may have a large impact on the prices.
of long-dated path-dependent options, even when they have no explicit hybrid features, and thus should not be ignored when pricing such contracts.

Appendix 5.A  Proofs

Proof of theorem 5.2.1. It was shown in the appendix to chapter 3 that the short-rate under the Hull-White model satisfies

\[ r_t = \tilde{r}_t + \xi_t, \]

where

\[ \tilde{r}_t = \int_0^t h(s, t) dW_{2,t}, \quad \xi_t = f(0, t) + \int_0^t h(s, t) H(s, t) ds, \]

\[ h(u, v) := \psi_u e^{-u a x z} - f u \int_a^z dw, \quad H(u, t) := t \int_u^t h(u, v) dv. \]

Now, omitting the subscript \( k \) for clarity, define

\[ v_t := \eta_t^2 \]

and

\[ \tilde{x}_t := \log(S_t) + D_t - \Upsilon_t, \quad D_t := \int_0^t y_u du, \quad \Upsilon_t := \int_0^t \xi_u du. \]

Then \( Y_t := (\tilde{x}_t, \tilde{r}_t, v_t, \eta_t) \) satisfies the following affine system of SDEs.

\[ d\tilde{x}_t = \left( \tilde{r}_t - m_t^2 v_t^2/2 \right) dt + m_t \eta_t d\tilde{W}_{1,t}, \]

\[ d\tilde{r}_t = -a_t \tilde{r}_t dt + \psi_t d\tilde{W}_{2,t}, \]

\[ dv_t = \left( 2\kappa_t \eta_t - 2\kappa_t v_t + \gamma_t^2 \right) dt + 2\gamma_t \eta_t d\tilde{W}_{3,t}, \]

\[ d\eta_t = \kappa_t (\bar{\eta}_t - \eta_t) dt + \gamma_t d\tilde{W}_{3,t}. \]

We now derive the \( T \)-forward characteristic function of \( Y_t \) by applying the results of Duffie et al. (2000). Using the notation of their paper, we begin by writing

\[ dY_t = \hat{\mu}(Y_t, t) dt + \hat{\Sigma}(Y_t, t) dW_t, \]

where \( \hat{\mu} : \mathbb{R}^4 \times [0, \infty) \rightarrow \mathbb{R}^4 \) and \( \hat{\Sigma} : \mathbb{R}^4 \times [0, \infty) \rightarrow \mathbb{R}^4 \) have the forms

\[ \hat{\mu}(x, t) = K_0(t) + K_1(t)x, \quad \hat{\Sigma x} = H_0(t) + \sum_{k=1}^4 H_{1,k}(t)x_k, \]

with time-dependent coefficients

\[ K_0(t) = (0, 0, \gamma_t^2, \kappa_t \bar{\eta}_t), \]
\[ K_1(t) = \begin{pmatrix} 0 & 1 & -\frac{1}{2} m_t^2 & 0 \\ 0 & -a_t & 0 & 0 \\ 0 & 0 & -2\kappa_t & 2\kappa_t \bar{\eta}_t \\ 0 & 0 & 0 & -\kappa_t \end{pmatrix}, \]

\[ H_0(t) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \psi_t^2 & 0 & \psi_t \gamma t \rho_{t,2,3} \\ 0 & 0 & 0 & 0 \\ 0 & \psi_t \gamma t \rho_{t,2,3} & 0 & \gamma_t^2 \end{pmatrix}, \]

\[ H_{1,1}(t) = H_{1,2}(t) = 0, \]

\[ H_{1,3}(t) = \begin{pmatrix} m_t^2 & 0 & 2\gamma_t m_t \rho_{t,1,3} & 0 \\ 0 & 0 & 0 & 0 \\ 2\gamma_t m_t \rho_{t,1,3} & 0 & 4\gamma_t^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \]

\[ H_{1,4}(t) = \begin{pmatrix} 0 & \psi_t m_t \rho_{t,1,2} & 0 & \gamma_t m_t \rho_{t,1,3} \\ \psi_t m_t \rho_{t,1,2} & 0 & 2\psi_t \gamma t \rho_{t,2,3} & 0 \\ 0 & 2\psi_t \gamma t \rho_{t,2,3} & 0 & 2\gamma_t^2 \\ \gamma_t m_t \rho_{t,1,3} & 0 & 2\gamma_t^2 & 0 \end{pmatrix}. \]

Also, continuing to follow the notation of Duffie et al. (2000), let \( R : \mathbb{R}^4 \times [0, \infty) \to \mathbb{R} \) be a discount rate function of the form

\[ R(x, t) = \varrho_0(t) + \varrho_1(t) \cdot x, \]

with time-dependent coefficients \( \varrho_0(t) \in \mathbb{R} \) and \( \varrho_1(t) \in \mathbb{R}^4 \). Then, using equations (2.3) to (2.6) of their paper, we have that the discounted characteristic function of \( Y_T \), for \( u := (u_1, u_2, u_3, u_4) \in \mathbb{C}^4 \) and \( t \in [0, T] \), is

\[ \tilde{\phi}(u, Y_t, t, T) = \mathbb{E} \left( e^{-\int_t^T R(Y_{s,s})ds + iu \cdot Y_T} \right)_{\mathcal{F}_t} = e^{\alpha(t) + \beta(t) \cdot Y_t}, \]

where \( \alpha : [0, T] \to \mathbb{C} \) and \( \beta : [0, T] \to \mathbb{C}^4 \) satisfy the differential equations

\[ \frac{d}{dt} \beta(t) = \varrho_1(t) - K_1^r(t) \beta(t) - \frac{1}{2} \sum_{k=0}^{4} (\beta^r(t) H_{1,k}(t) \beta(t)) e_k, \]

\[ \frac{d}{dt} \alpha(t) = \varrho_0(t) - K_0(t) \cdot \beta(t) - \frac{1}{2} \beta^r(t) H_0(t) \beta(t), \]

with terminal conditions \( \beta(T) = iu \) and \( \alpha(T) = 0 \). By defining \( B(\tau) = \beta(T - \tau) \) and \( A(\tau) = \alpha(T - \tau) \),
we can rewrite these equations as

\[ \frac{d}{d\tau} B(\tau) = -\varphi_1(T - \tau) + K_{1r}(T - \tau)B(\tau) + \frac{1}{2} \sum_{k=0}^{4} (B^{(r)}(\tau)H_{1,k}(T - \tau)B(\tau)) e_k, \]

\[ \frac{d}{d\tau} A(\tau) = -\varphi_0(T - \tau) + K_0(T - \tau) \cdot B(\tau) + \frac{1}{2} B^{(r)}(\tau)H_0(T - \tau)B(\tau), \]

for \( \tau \in [0, T] \), with initial conditions \( B(0) = iu \) and \( A(0) = 0 \). Setting \( R(x, t) = x_2 \), and assuming that all the time-dependent parameters are piecewise-constant, this translates into the following system of differential equations for \( \tau \in (\tau_j, \tau_{j+1}] \),

\[ \frac{d}{d\tau} B_1 = 0, \quad \text{(5.A.1)} \]

\[ \frac{d}{d\tau} B_2 = -1 + B_1 - \bar{a}_j B_2, \quad \text{(5.A.2)} \]

\[ \frac{d}{d\tau} B_3 = \frac{1}{2} \bar{m}_j^2 B_1 (B_1 - 1) + 2 (\bar{\gamma}_j \bar{m}_j \bar{\rho}_{j,1,3} B_1 - \bar{\kappa}_j) B_3 + 2 \bar{\gamma}_j \bar{B}_3^2, \quad \text{(5.A.3)} \]

\[ \frac{d}{d\tau} B_4 = 2 \bar{\kappa}_j \bar{\eta}_j B_3 + \bar{\psi}_j \bar{m}_j \bar{\rho}_{j,1,2} B_1 B_2 + 2 \bar{\psi}_j \bar{\gamma}_j \bar{\rho}_{j,2,3} B_2 B_3 + \left( \bar{\gamma}_j \bar{m}_j \bar{\rho}_{j,1,3} B_1 - \bar{\kappa}_j + 2 \bar{\gamma}_j^2 B_3 \right) B_4, \quad \text{(5.A.4)} \]

\[ \frac{d}{d\tau} A = \bar{\gamma}_j^2 B_3 + \frac{1}{2} \bar{\psi}_j^2 B_2^2 + \left( \bar{\kappa}_j \bar{\eta}_j + \bar{\psi}_j \bar{\gamma}_j \bar{\rho}_{j,2,3} B_2 + \frac{1}{2} \bar{\gamma}_j^2 B_4 \right) B_4, \quad \text{(5.A.5)} \]

where the \( \tau \) dependence in \( A \) and \( B \) := \( B_1, B_2, B_3, B_4 \) has been omitted to ease the notation. Define \( B_{i,j} := B_i(\tau_j) \) and \( A_j := A(\tau_j) \). Solving ODEs (5.A.1) and (5.A.2) yields

\[ B_1 = iu_1, \quad B_2 = \frac{iu_1 - 1}{a_j} + \left( B_{2,j} - \frac{iu_1 - 1}{a_j} \right) e^{-\bar{a}_j(\tau - \tau_j)}. \]

ODE (5.A.3) is a Riccati equation with constant coefficients

\[ b_0 := -\frac{1}{2} \bar{m}_j^2 u_1 (u_1 + i), \quad b_1 := 2 (\bar{\gamma}_j \bar{m}_j \bar{\rho}_{j,1,3} i u_1 - \bar{\kappa}_j), \quad b_2 := 2 \bar{\gamma}_j^2. \]

Therefore, as shown in appendix A of Wu and Zhang (2006), the solution is

\[ B_3 = B_{3,j} + \left( \frac{1 - e^{d(\tau - \tau_j)}}{1 - ge^{d(\tau - \tau_j)}} \right) h, \]

where

\[ d := \sqrt{b_1^2 + 4b_0 b_2}, \quad g := \frac{-b_1 + d - 2b_2 B_{3,j}}{-b_1 - d - 2b_2 B_{3,j}}, \quad h := \frac{-b_1 + d - 2b_2 B_{3,j}}{2b_2}. \]
To solve the ODE (5.A.4), first define the time-dependent coefficients
\[ q_0 (\tau) := 2\tilde{\kappa}_j \tilde{\eta}_j B_3 + \psi_j \tilde{m}_j \tilde{\rho}_{j,1,2} B_1 B_2 + 2\tilde{\psi}_j \tilde{\gamma}_j \tilde{\rho}_{j,2,3} B_2 B_3, \quad q_1 (\tau) := \tilde{\gamma}_j \tilde{m}_j \tilde{\rho}_{j,1,3} B_1 - \tilde{\kappa}_j + 2\tilde{\gamma}^2_j B_3. \]

Then the solution is
\[
B_4 = e^{\tau_j} \left( B_{4,j} + \int_{\tau_j}^{\tau} q_0 (y) e^{-\int_{\tau_j}^{y} q_1 (s) ds} dy \right),
\]
where
\[
\int_{\tau_j}^{\tau} q_1 (s) ds = \frac{d}{2} (\tau - \tau_k) - \log \left( \frac{1 - g e^{d(\tau - \tau_j)}}{1 - g} \right), \quad \int_{\tau_j}^{\tau} q_0 (y) e^{-\int_{\tau_j}^{y} q_1 (s) ds} dy = f_5 (\tau),
\]
and the function \( f_5 \) is as defined in theorem 5.2.1. Next, solving ODE (5.A.5) yields
\[
A = A_j + \int_{\tau_j}^{\tau} \left( \gamma_j^2 B_3 + \frac{1}{2} \tilde{\psi}_j^2 B_2^2 \right) dy + \int_{\tau_j}^{\tau} \left( \tilde{\kappa}_j \tilde{\eta}_j + \psi_j \tilde{\gamma}_j \tilde{\rho}_{j,2,3} B_2 + \frac{1}{2} \tilde{\gamma}^2_j B_4 \right) B_4 dy.
\]

An analytical expression for the first integral is given in equation (5.2.1), whereas the second integral involves the ordinary hypergeometric function, \( _2F_1 \), and must be computed numerically, as explained in section 5.2.

Now that we have an expression for the discounted characteristic function, \( \hat{\phi} \), we can determine the \( T \)-forward characteristic function (at time zero) as follows:
\[
\hat{\phi}_Y (u) = E^T \left( e^{iuY_T} \right) = P(0, T)^{-1} E \left( -\int_0^T r_s ds + iuY_T \right)
\]
\[
= P(0, T)^{-1} E \left( -\int_0^T \tilde{r}_s ds + iuY_T + \int_0^T \tilde{\xi}_s ds + A(T) + B(T) \cdot Y_0 \right).
\]
Furthermore, the joint PDF of \( \tilde{X}_T := (\tilde{\xi}_T, \tilde{r}_T, \tilde{\eta}_T) \) in the \( T \)-forward measure is equal to the Fourier transform of \( \hat{\phi}_X (w) = \hat{\phi}_Y (w_1, w_2, 0, w_3) \),
\[
\phi_X (\tilde{x}) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} e^{-iw \cdot \tilde{x}} \hat{\phi}_X (w) \, dw,
\]
where $\tilde{x} := (\tilde{x}_1, \tilde{x}_2, \tilde{x}_3)$ and $w := (w_1, w_2, w_3)$. Next, in order to determine the joint PDF of $X_T := (S_T, r_T, \eta_T)$, define the bijective function $h : \mathbb{R}^+ \times \mathbb{R}^2 \rightarrow \mathbb{R}^3$, such that

$$h_1(x) = \log(x_1) + D_T - \Upsilon_T, \quad h_2(x) = x_2 - \xi_T, \quad h_3(x) = x_3,$$

where $h := (h_1, h_2, h_3)$ and $x := (x_1, x_2, x_3)$. Therefore the joint PDF of $X_T = h^{-1} (\tilde{X}_T)$ is

$$\phi_X(x) = |\det (J_h)| \phi_{\tilde{X}}(h(x)),$$

where the Jacobian matrix is

$$J_h = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$
Chapter 6

Non-Parametric Pricing of Volatility Derivatives under Stochastic Interest Rates

6.1 Introduction

As observed by Carr and Lee (2009), variance swaps and other more complex volatility derivatives are increasingly being used by organisations to either trade volatility or hedge their portfolio's vega exposure. However, much of the existing literature is either heavily model dependent, only applies to the standard variance swap, or assumes that interest rates are deterministic. Although this last assumption may be sufficient for short-term contracts, the bias due to stochastic rates could be quite significant for long-term ones. Furthermore, as observed by Rebonato (2004), a vanilla option's implied volatility is in fact the volatility of the forward price, not the spot price, and thus includes a component due to bond price volatility. We therefore study the impact of stochastic interest rates on volatility derivatives pricing by extending the non-parametric approach of Carr and Lee (2008).

With only a few assumptions on the underlying stock's dynamics, Carr and Lee (2008) proved that the prices of general volatility derivatives, not just variance swaps, are determined given the complete continuum of calls and puts for all strikes. They began by relating the price of a power option, whose payoff at expiry depends on the stock price raised to a given power, to that of an exponential stock variance contract. They then determined the price of various classes of volatility derivatives by relating the payoff function back to the exponential function. Their two key assump-
tions were that the asset price process is jump free, and that the volatility process and driving Brownian motion are independent. The effect of this second assumption was mitigated by reducing the sensitivity of the resulting price to correlation between the processes. However, like much work on volatility derivatives pricing, interest rates were assumed to be zero. Although the extension to time-dependent deterministic rates is straightforward, allowance for stochastic rates is not.

We extend the above approach by allowing the bond price to follow a diffusive process that satisfies pull to par and certain independence assumptions. We then show that the price of a power option is equal to that of a exponential variance contract, whose payoff depends on the sum of the quadratic variations of the log price processes of the stock and the bond, under additional independence assumptions. Next, given a model for interest rates, we can isolate the value of an exponential stock variance contract. Thus the Laplace transform of the density of stock variance is determined by the expected exponential bond variance and the price of a non-path-dependent power option. This result is semi-model-free in the sense that we only require an explicit parametrization of the interest rate processes (in order to calculate the expected exponential bond variance), and not of the stock price process itself. Furthermore, it may be possible to determine the expected exponential bond variance from the prices of interest rate derivatives in a non-parametric way. However, it is not clear that there is enough market information to do so in practice, and we assume a short rate model when giving empirical examples later in the chapter. Also note that we do not address the issue of replication, as was done in Carr and Lee (2008), due to the complications introduced by interest rate risk.

A key assumption above is that the volatility process is independent of the stock’s driving Brownian motion. Therefore, it cannot explicitly depend on the stock price, as it would, for example, in a local volatility model. As explained by Gatheral (2006), this assumption also means that the Black-Scholes implied volatility smile, which refers to the dependence of implied volatility on strike, is symmetric about the forward price. However, smiles observed in the market do not typically satisfy this condition, implying that the zero correlation assumption is unrealistic. Fortunately, we can also extend the Carr and Lee (2008) “correlation neutral” approach to the case of non-deterministic interest rates. To this end, we write the value of an exponential variance contract in terms of a linear combination of two different power options, whose price is insensitive to the level of correlation, in the sense that its first derivative with respect to correlation is zero. Thus we achieve the same level of correlation neutrality as Carr and Lee.

Building on this, we show how to recover the price of a range of volatility derivatives given our expression for exponential variance contracts. For certain contracts, including variance swaps and other variance power payoffs under specific conditions, we are able to equate the price to that of a single non-path-dependent option, which we then approximate using a combination of calls/puts whose market prices are available. If this is not possible, we instead derive an expression involving an integral of power option prices, which is still determined by the observed market smile.

Next, in order to give empirical examples based on real market data, we show how to modify the mixture of normals method of Rebonato and Cardoso (2004) to explicitly allow for independent
6.1. Introduction

stochastic interest rates when fitting a density for the final stock price in the T-forward measure to observed vanilla option prices. This density is then adjusted to remove the effect of interest rate stochasticity while holding the T-expiry zero-coupon bond price constant. This procedure does not affect the price of volatility derivatives, and therefore provides a practical way to apply existing results for the pricing volatility derivatives to the case of independent stochastic rates.

A number of other authors have focused on the model-free valuation of volatility derivatives. One of the earliest results in this area is due to Neuberger (1994). He showed that a variance swap can be replicated using a static position in log contracts and continuous trading of the underlying, assuming only that the underlying follows a continuous process whose sum of squared deviations approach a finite limit. Although Neuberger suggested that the log contract should be traded to enable his replicating strategy, Breeden and Litzenberger (1978) established that any European payoff profile, including the log price, can be replicated given the continuum of call prices for all strikes and maturities, under certain integrability conditions.

Following this work, Derman, Kamal, Kani, and Zou (1996) demonstrated how variance swaps, and certain other contracts whose payoffs depend on realized variance, can be priced and hedged using only vanilla European options. Furthermore, Carr and Madan (1998) explicitly detailed the profit obtained by delta hedging a static position in a given option, and pointed out that if the option in question is two times a log contract then the profit is equal to the payoff of a variance swap. They went on to show how the same method can be used to replicate a variety of other volatility contracts. As with previous model-free work, they assumed that the price process is continuous and that interest rates are constant.

More recently, Shen and Siu (2013) have studied the price of variance swaps under stochastic interest rates. They specifically focused on a regime-switching Schöbel-Zhu model for volatility, coupled with a regime-switching Hull-White model for interest rates. They developed an integral expression for the price of a variance swap under this model, and investigated the effect of the parameters on the price. Conversely, Hörfelt and Torné (2010) examined variance swaps in a non-parametric setting, assuming that the bond and the stock follow diffusive processes. However, their work focused purely on the standard variance swap, and is not applicable to a wide variety of volatility derivatives found in the market, including volatility swaps (whose payoff depends on the square-root of observed variance) or call/puts on variance.

In section 6.2 below, we introduce the notation and assumptions that will apply throughout this chapter. Next, in section 6.3, we show that the price of an exponential variance contract is equal to that of a non-path-dependent power option. Then, in section 6.4, we establish that the Carr and Lee (2008) concept of correlation neutrality can be extended to independent stochastic interest rates. This culminates with the result that the correlation neutral price of an exponential stock variance contract is determined by the price of a exponential bond variance contract and a linear combination of two different power options. Using this result, we explain how to recover the prices of a variety of other volatility derivatives, including variance swaps and volatility swaps, in section 6.5. Following this, in section 6.6, we extend our results to processes with unbounded quadratic vari-
ation. Then, given that call and put option prices are only available for a finite number of strikes, we demonstrate, in section 6.7, how a mixture of normals can be used to fit the market smile under stochastic interest rates, which in turn allows us to apply the results of the previous sections. Using this approach, we give some empirical examples of the effects of non-deterministic interest rates on long-dated volatility derivatives in section 6.8. Finally, we conclude in section 6.9.

6.2 Notation and Assumptions

Consider a filtered probability space \((\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)\) satisfying the usual conditions. Let \(r_t, \eta_t\) and \(\psi_t\) be some measurable \(\mathcal{F}_t\)-adapted processes, and fix an expiry time, \(T > 0\). Our first assumption specifies the form of the processes for the stock price and bond price.

**Assumption 6.2.1.** We assume that markets are frictionless and that there exists an equivalent risk-neutral measure, \(Q\), such that the stock price, \(S_t\), and \(T\)-expiry zero-coupon bond price, \(P_t := P(t, T)\), follow diffusive processes of the form

\[
\begin{align*}
dS_t &= r_t S_t dt + \eta_t S_t dW^S_t, \\
dP_t &= r_t P_t dt + \psi_t P_t dW^P_t,
\end{align*}
\]

where \((W^S, W^P)\) is a two-dimensional \(\mathcal{F}_t\)-adapted \(Q\)-Brownian motion with constant correlation \(\rho\). We also assume that the process for \(P_t\) satisfies pull to par, i.e. \(P(T, T) = 1\).

Here \(\eta_t\) and \(\psi_t\) are the volatility processes of the stock price and bond price respectively, and \(r_t\) is the instantaneous interest rate. Furthermore, as \(Q\) is the risk-neutral measure, the no-arbitrage prices of derivatives are equal to their expected discounted payoffs under this measure.

In order for \(P_t\) to satisfy pull to par, the stochastic processes \(W^P_t\), \(\psi_t\) and \(r_t\) must be intimately related. Thus, when constructing a model for the bond price, we will first specify a process for \(r_t\), and then determine the implied process for \(\psi_t\) by applying Ito’s lemma to the bond pricing equation,

\[
P_t = E \left( e^{-\int_t^T r_s ds} \bigg| \mathcal{F}_t \right).
\]

Although we leave \(r_t\) and \(\psi_t\) unspecified here, in section 6.8 we will assume that \(r_t\) follows the Hull-White model. This model, along with many other short rate models, admits an affine term structure (Björk, 2009, proposition 24.8), which means that bond prices take the form

\[
P_t = e^{A(t, T) - B(t, T) r_t}, \quad \forall \ t \in [0, T],
\]

where \(A(t, T)\) and \(B(t, T)\) are deterministic functions such that \(A(T, T) = B(T, T) = 0\). Given a short rate model with risk-neutral dynamics \(dr_t = \mu(t, r_t) dt + \sigma(t, r_t) dW_t^r\), applying Ito’s lemma to
equation (6.2.1) yields
\[
\frac{dP_t}{P_t} = \left( \frac{\partial A(t, T)}{\partial t} - r_t \frac{\partial B(t, T)}{\partial t} - \mu(t, r_t) B(t, T) + \frac{1}{2} \sigma^2(t, r_t) B^2(t, T) \right) dt - \sigma(t, r_t) B(t, T) dW^r_t.
\]

In an arbitrage-free model, the drift coefficient above must equal the short rate (Björk, 2009, equation 24.17). Thus, all arbitrage-free affine term structure models satisfy assumption 6.2.1 with \( \psi_t = \sigma(t, r_t) B(t, T) \). In particular, under the Hull-White model, \( \sigma(t, r_t) \) is a constant, which implies that \( \psi_t \) is a deterministic function of \( t \) and \( T \) only. Indeed, any affine term structure model that has a deterministic \( \sigma(t, r_t) \) will also have a deterministic \( \psi_t \).

We next assume that the volatility processes, \( \eta_t \) and \( \psi_t \), are independent of the driving Brownian motions, \( W^S_t \) and \( W^P_t \), i.e.

**Assumption 6.2.2.**

\( (\eta, \psi) \perp (W^S, W^P) \).

We initially allow \( W^S \) and \( W^P \) to have constant correlation, \( \rho \), in order to derive a more general expression for the value of a power option, which pays \( (S_T/F_0)^p \) at time \( T \). This leads to equation (6.3.2) in section 6.3, which equates the price of a power option to that of an exponential variance contract that depends on \( \eta, \psi \) and \( \rho \). Then, in order to separate the terms depending on \( \eta \) and \( \psi \), we assume that \( \rho = 0 \), and that \( \eta \) and \( \psi \) are independent, i.e.

**Assumption 6.2.3.**

\( W^S \perp W^P \) and \( \eta \perp \psi \).

Let \( X_t \) be the log stock price process, and \( \langle X \rangle_t \) be its quadratic variation. Likewise, let \( Y_t \) be the log bond price process, and \( \langle Y \rangle_t \) be its quadratic variation, i.e.

\[
X_t := \log(S_t), \quad Y_t := \log(P_t),
\]

\[
\langle X \rangle_t = \int_0^t \eta_u^2 du, \quad \langle Y \rangle_t = \int_0^t \psi_u^2 du.
\]

We make the additional assumption that these quadratic variations are bounded, i.e.

**Assumption 6.2.4.**

\( \langle X \rangle_T \leq m_1 \) and \( \langle Y \rangle_T \leq m_2 \) for some non-negative reals \( m_1 \) and \( m_2 \).

Of course, many popular stochastic volatility models, such as the Heston model, do not have bounded quadratic variation. However, under deterministic interest rates, Carr and Lee (2008) have explained how the assumption that \( \langle X \rangle_T \leq m_1 \) can be dropped, in the sense that, given certain additional conditions, their pricing formulas are still valid without it. We have extended this result
to include stochastic interest rates, so that assumption 6.2.4 can similarly be discarded. For details of the proof, and on the required conditions, see section 6.6.

Throughout this chapter we will focus on the valuation of volatility derivatives at time zero. However, as our results apply to general volatility derivatives, including those that are seasoned at time zero, the valuation of contracts at times greater than zero is straightforward. When we discuss such cases, we will use \( \mathbb{E}_T^T (\cdot) := \mathbb{E}^T (\cdot | F_t) \) to denote the expectation in the \( T \)-forward measure, conditional on the information up to time \( t \). Also, we will use \( F_t := S_t / P_t \) to denote the \( T \)-forward price at time \( t \).

Note that our results do not rely on the interest rate process following an affine term structure model, and that we only use the Hull-White model to compute the value of \( \mathbb{E}^T (e^{\lambda \langle Y \rangle_T}) \), which appears in theorem 6.3.1. The complete set of restrictions that we place on the interest rate model are given in assumptions 6.2.1 to 6.2.4.

### 6.3 Exponential Variance Contracts

Our first goal is to relate the price of a power option to that of an exponential stock variance contract, which pays \( e^{\lambda \langle X \rangle_T} \) at time \( T \). We begin by changing from the risk-neutral measure to the \( T \)-forward measure. The solution to the SDE for the bond price is

\[
P_t = P_0 e^{\int_0^t r_u du} - \frac{1}{2} \int_0^t \psi_u^2 du + \int_0^t \psi_u dW^P_u.
\]

If we let the bank account be \( \beta_t := e^{\int_0^t r_u du} \), then the Radon-Nikodym derivative of the \( T \)-forward measure, \( \mathbb{Q}^T \), with respect to the risk-neutral measure, \( \mathbb{Q} \), is

\[
\frac{d\mathbb{Q}^T}{d\mathbb{Q}} \bigg|_{F_t} = \frac{P_t}{P_0 \beta_t} = e^{-\frac{1}{2} \int_0^t \psi^2_u du + \int_0^t \psi_u dW^P_u}.
\]

Therefore, by the Girsanov theorem, we have

\[
dW^{P,T}_t = -\psi_t dt + dW^P_t, \\
dW^{S,T}_t = -\rho \psi_t dt + dW^S_t,
\]

where \( W^{P,T}_t \) and \( W^{S,T}_t \) are Brownian motions in the \( T \)-forward measure. Solving the resulting SDEs for \( X_T \) and \( Y_T \) yields

\[
X_T = X_0 + \int_0^T r_u du + \int_0^T \rho \psi_u \eta_u du - \frac{1}{2} \int_0^T \psi_u^2 du + \int_0^T \eta_u dW^{S,T}_u, \quad \text{and}
\]

\[
Y_T = Y_0 + \int_0^T r_u du + \frac{1}{2} \int_0^T \psi_u^2 du + \int_0^T \psi_u dW^{P,T}_u.
\]
Combining these two equations, and using the fact that \( Y_T = \log(P(T, T)) = 0 \), we have the following expression for the log stock price, \( X_T \), in the T-forward measure:

\[
X_T = X_0 - Y_0 - \frac{1}{2} \langle Y \rangle_T - \int_0^T \psi_udW_u^{P,T} + \int_0^T \rho \psi_u \eta_u du - \frac{1}{2} \langle X \rangle_T + \int_0^T \eta_u dW_u^{S,T} .
\] (6.3.1)

So, conditional on the stock and bond volatility paths, \( H_T := \sigma (F_T, F_T^\psi) \), \( X_T - X_0 + Y_0 \) has a normal distribution with mean and variance as follows:

\[
E_T (X_T - X_0 + Y_0 | H_T) = -\frac{1}{2} \left( \langle Y \rangle_T - 2 \int_0^T \rho \psi_u \eta_u du + \langle X \rangle_T \right),
\]

and

\[
\text{Var}_T (X_T - X_0 + Y_0 | H_T) = \langle Y \rangle_T - 2 \int_0^T \rho \psi_u \eta_u du + \langle X \rangle_T .
\]

For \( p \in \mathbb{C} \), let \( D_p \) denote the power option that pays \( (S_T / F_0)^p \) at time \( T \). Its value is

\[
D_p(0) = P_0 \mathbb{E}_T \left( \left( \frac{S_T}{F_0} \right)^p \right) = P_0 \mathbb{E}_T \left( e^{p(X_T - X_0 + Y_0)} | H_T \right)) = P_0 \mathbb{E}_T \left( e^{\frac{1}{2}p(p-1)(\langle Y \rangle_T - 2 \int_0^T \rho \psi_u \eta_u du + \langle X \rangle_T)} \right) .
\] (6.3.2)

Now assuming that \( \rho = 0 \), and that \( \psi \) and \( \eta \) are independent, we can write the price of an exponential stock variance contract, which we call \( H_\lambda \), purely in terms of the price of a power option and an exponential bond variance contract:

\[
H_\lambda(0) = P_0 \mathbb{E}_T \left( e^{\lambda \langle X \rangle_T} \right) = P_0 \mathbb{E}_T \left( e^{\lambda \langle Y \rangle_T} \right)^{-1} \mathbb{E}_T \left( \left( \frac{S_T}{F_0} \right)^{p_+} \right) ,
\]

where \( p_+ = \frac{1}{2} + \sqrt{\frac{1}{4} + 2 \lambda} \), \( p_- = \frac{1}{2} - \sqrt{\frac{1}{4} + 2 \lambda} \), and \( \lambda \in \mathbb{C} \). Note that we have the freedom to take a linear combination of the given power options, which leads us to theorem 6.3.1 below. The coefficients \( g_1 \) and \( g_2 \) in this linear combination will be chosen in section 6.4 so that the resulting expression is first-order correlation neutral.

**Theorem 6.3.1.** Let \( p_\pm \) be defined as above, and let \( g_1, g_2 \in \mathbb{R} \) such that \( g_1 + g_2 = 1 \). Under assumptions 6.2.1 to 6.2.3, the price of an exponential stock variance contract with parameter \( \lambda \) is

\[
H_\lambda(0) = P_0 \mathbb{E}_T \left( e^{\lambda \langle Y \rangle_T} \right)^{-1} \mathbb{E}_T \left( \left( \frac{S_T}{F_0} \right)^{p_+} g_1 + \left( \frac{S_T}{F_0} \right)^{p_-} g_2 \right) .
\]

This theorem implies that, given a model for interest rates, the Laplace transform of the density of quadratic variation is determined by the complete continuum of vanilla option prices. Thus the density itself, and the prices of general volatility derivatives, are also determined. Furthermore,
the same way that vanilla call prices determine the density of the stock price at expiry, the complete continuum of variance call prices determines the density of quadratic variation at expiry. This in turn determines the prices of vanilla options, and hence the relationship between vanilla options and variance options works both ways under our assumptions.

In the remainder of this chapter we look at various ways to recover the prices of volatility derivatives from the observed market prices of calls and put. However, to address the potential for non-zero correlation between the volatility process and the stock’s driving Brownian motion, we first examine the concept of correlation neutrality.

### 6.4 Correlation Neutrality

A key assumption of the previous section is that the stock’s volatility, \( \eta \), and the driving Brownian motion, \( W^S \), are independent. However Black (1975) observed that downwards price shocks are linked to greater volatility than comparable upwards shocks to US equity indices. This implies that the correlation, \( \rho_\eta \), between \( \eta \) and \( W^S \) is negative, with Lewis (2000) having suggested a range of \(-0.5\) to \(-0.8\). In order to mitigate the effect of this correlation, Carr and Lee (2008) demonstrated that volatility derivatives can be valued using non-path-dependent options whose prices have first derivative with respect to \( \rho_\eta \) of zero when evaluated at \( \rho_\eta = 0 \). Below, we will extend this concept, called “correlation neutrality” or “correlation immunity”, to independent stochastic interest rates.

The first step is to derive an expression for the value of a European option in terms of \( \rho_\eta \). To this end, let \( (W_{1,t}, W_{2,t}) \) be an \( F_t \)-adapted two-dimensional standard Brownian motion under \( Q^T \). The following assumption specifies how the volatility process, \( \eta \), and driving Brownian motion, \( W^{S,T} \), are correlated:

**Assumption 6.4.1.** We assume that \( \eta_t, \psi_t \) and \( W_{2,t} \) are adapted to some filtration \( (\mathcal{G}_t)_{t \geq 0} \) such that \( \mathcal{G}_t \subseteq \mathcal{F}_t \forall t \geq 0 \), and

\[
dW^{S,T}_t = \sqrt{1 - \rho_\eta^2} dW_{1,t} + \rho_\eta dW_{2,t}, \quad \rho_\eta \in (-1, 1).
\]

We note that this set up allows \( \eta \) to be a discontinuous process, as long as the jumps are independent of \( W_1 \) and \( W^{P,T} \). By substituting the above expression for \( dW^{S,T}_t \) into equation (6.3.1), setting \( \rho = 0 \), and then conditioning on \( \mathcal{G}_T \), we see that \( X_T|\mathcal{G}_t \) has a normal distribution with mean and variance as follows:

\[
E^T(X_T|\mathcal{G}_t) = X_0 - Y_0 - \frac{1}{2} (\langle Y \rangle_T + \langle X \rangle_T) + \rho_\eta \int_0^T \eta_u dW_{2,u}, \quad \text{and}
\]

\[
\text{Var}^T(X_T|\mathcal{G}_t) = \langle Y \rangle_T + (1 - \rho_\eta^2) \langle X \rangle_T.
\]

Therefore the undiscounted price of a European option, paying \( F(S_T) \) at time \( T \), is equal to the
Theorem 6.4.1. Let \( v \) be given by the following theorem:

\[
E^T (F(S_T)) = E^T (E^T (F(e^{X_T}) | G_T))
\]

\[
= E^T \left( \int_0^\infty F \left( ye^{\rho_\eta \int_0^T \eta_u dW_{2,u}} \right) \phi_{\mu_1,v_1}(y) dy \right),
\]

where \( v_1 := \text{Var}^T (X_T|G_t), \mu_1 := \text{E}^T (X_T|G_t) = \rho_\eta \int_0^T \eta_u dW_{2,u} \) and \( \phi_{\mu,v} \) is the log-normal \((\mu,v)\) density. By evaluating the right hand side of the above equation at \( \rho_\eta = 0 \), and its first derivative with respect to \( \rho_\eta \) at \( \rho_\eta = 0 \), we can write down the first-order Maclaurin expansion of the undiscounted option price with respect to \( \rho_\eta \):

\[
E^T (F(S_T)) = E^T \left( \int_0^\infty F(y) \phi_{\mu_1,v_2}(y) dy \right) + \frac{\partial}{\partial \rho_\eta} E^T \left( \int_0^\infty yF'(y) \phi_{\mu_1,v_2}(y) dy \int_0^T \eta_u dW_{2,u} \right) \rho_\eta + O(\rho_\eta^2),
\]

where \( v_2 := \text{Var}^T (X_T|\rho_\eta = 0, G_t) = \langle Y \rangle_T + \langle X \rangle_T. \) If the coefficient of \( \rho_\eta \) in the above expansion is equal to zero, then we say that the option is first-order correlation neutral. A sufficient condition for this is given by the following theorem:

**Theorem 6.4.1.** Let \( \mu_1 = X_0 - Y_0 - \frac{1}{2} \left( \langle Y \rangle_T + \langle X \rangle_T \right) \) and \( v_2 = \langle Y \rangle_T + \langle X \rangle_T. \) Under assumptions 6.2.1 and 6.4.1, if a European option's payoff function, \( F(S_T) \), satisfies

\[
\int_0^\infty yF'(y) \phi_{\mu_1,v_2}(y) dy = c,
\]

where \( c \) does not depend on \( \langle Y \rangle_T \) or \( \langle X \rangle_T \), then the first derivative of its price with respect to \( \rho_\eta \) will be 0 when evaluated at \( \rho_\eta = 0 \). We call such options “first-order correlation neutral”.

**Proof.** Suppose \( \int_0^\infty yF'(y) \phi_{\mu_1,v_2}(y) dy = c \), where \( c \) does not depend on \( \langle Y \rangle_T \) or \( \langle X \rangle_T \). When evaluated at \( \rho_\eta = 0 \), the first derivative of \( E^T (F(S_T)) \) with respect to \( \rho_\eta \) is

\[
\frac{\partial}{\partial \rho_\eta} E^T (F(S_T)) \bigg|_{\rho_\eta=0} = \frac{\partial}{\partial \rho_\eta} E^T \left( \int_0^\infty F \left( ye^{\rho_\eta \int_0^T \eta_u dW_{2,u}} \right) \phi_{\mu_1,v_1}(y) dy \right) \bigg|_{\rho_\eta=0}
\]

\[
= E^T \left( \int_0^\infty yF'(y) \phi_{\mu_1,v_2}(y) dy \int_0^T \eta_u dW_{2,u} \right)
\]

\[
= E^T \left( c \int_0^T \eta_u dW_{2,u} \right) = 0.
\]

\( \square \)

For example, any contract with an affine payoff function, \( F_{a,b}(S_T) := aS_T + b \), is first-order correlation neutral. In fact, the undiscounted price at time zero, \( E^T (F_{a,b}(S_T)) = aS_0/P_0 + b \), does
Let $\rho$ be the market bond price and option prices constant, the effect of assuming that interest rates do not depend on $\rho$ at all. Checking the condition in theorem 6.4.1, we see that

$$\int_0^\infty y F_a(b) \phi_{\mu_1,v_2}(y) \, dy = a \int_0^\infty y \phi_{\mu_1,v_2}(y) \, dy$$

$$= a e^{\mu_1 + \frac{1}{2} v_2}$$

$$= a e^{X_0 - Y_0},$$

which does not depend on $\langle Y \rangle_T$ or $\langle X \rangle_T$. We note that this condition on $F(y)$ is equivalent to that found in Carr and Lee (2008) under zero interest rates, which states that an option is correlation neutral if its delta under the Black-Scholes model is independent of the volatility parameter.

We can now use theorem 6.4.1 to choose $g_1$ in theorem 6.3.1 so that the European option used to price the exponential variance contract is correlation neutral:

**Theorem 6.4.2.** Let $D_{exp}$ be a European derivative that pays $F_{exp}(S_T; \lambda)$ at time $T$, where

$$F_{exp}(S_T; \lambda) := E^T \left( e^{\lambda(Y)_T} \right)^{-1} \left( \frac{S_T}{F_0} \right)^{p_+} g_1 \left( \frac{S_T}{F_0} \right)^{p_-} g_2,$$

$$p_+ = \frac{1}{2} + \frac{\sqrt{1 + 2\lambda}}{4}, \quad p_- = \frac{1}{2} - \frac{\sqrt{1 + 2\lambda}}{2},$$

$$g_1 = \frac{1}{2} - \frac{1}{2\sqrt{1 + 8\lambda}}, \quad g_2 = \frac{1}{2} + \frac{1}{2\sqrt{1 + 8\lambda}},$$

then $D_{exp}$ is first-order correlation neutral. Also, under assumptions 6.2.1, 6.2.2 and 6.2.3, the price of an exponential stock variance contract, which pays $e^{\lambda(X)_T}$ at time $T$, is equal to the price of $D_{exp}$, i.e.

$$E^T \left( e^{\lambda(X)_T} \right) = E^T \left( F_{exp}(S_T; \lambda) \right).$$

**Proof.** The final equality was proved in section 6.3. In order to prove that $D_{exp}$ is correlation neutral, first note that $g_1$ and $g_2$ solve the equation $g_1 p_+ + g_2 p_- = 0$, subject to the condition $g_1 + g_2 = 1$. Therefore, letting $F(y) = F_{exp}(y; \lambda)$, we have

$$\int_0^\infty y F'(y) \phi_{\mu_1,v_2}(y) \, dy = E^T \left( e^{\lambda(Y)_T} \right)^{-1} \int_0^\infty \left( \frac{y}{F_0} \right)^{p_+} g_1 \left( \frac{y}{F_0} \right)^{p_-} g_2 \phi_{\mu_1,v_2}(y) \, dy$$

$$= E^T \left( e^{\lambda(Y)_T} \right)^{-1} (g_1 p_+ + g_2 p_-) e^{\lambda(Y)_T} = 0.$$

Thus $D_{exp}$ is first-order correlation neutral by theorem 6.4.1. □

Comparing this result to that found in Carr and Lee (2008), we see that the key adjustments are multiplication by a factor of $E^T \left( e^{\lambda(Y)_T} \right)^{-1}$ and the inclusion of $F_0$ inside the two power payoffs. Note that the value of a European option paying $\left( \frac{S_T}{F_0} \right)^{p_+} g_1 + \left( \frac{S_T}{F_0} \right)^{p_-} g_2$ is determined by the market-observable zero-coupon bond price, $P_0$, and the complete continuum of call/put prices. Therefore, holding the market bond price and option prices constant, the effect of assuming that interest rates
are stochastic (i.e. bond volatility is non-zero) is encapsulated by the term $E^T (e^{\lambda Y_T})^{-1}$. It is this term that will cause complications when pricing other volatility derivatives using exponential variance contracts.

### 6.5 Other Variance Contracts

Given the (correlation neutral) price of exponential variance contracts, we can recover the price of a wide range of other volatility derivatives. Carr and Lee (2008) gave a variety of methods, each suited to a particular class of payoff functions, which we will extend to the case of stochastic interest rates. Where possible, we derive a result which equates the volatility derivative's price to that of a single European option, since this can then be approximated using market call and put prices. In other words, we derive equations of the form $E^T (h (\langle X \rangle_T)) = E^T (F_h (S_T))$, where $h$ is the volatility derivative's payoff function, and $F_h$ is the associated non-path-dependent option's payoff function.

We start by writing $h$ in terms of the exponential function. Typically, this means taking an affine function of $e^{\xi (z) \langle X \rangle_T}$, where $\xi$ is a given real or complex function of $z$, and then integrating or differentiating with respect to $z$. We then take the expectation of both sides in the forward measure, move the expectation through the integral or derivative and through the affine function, resulting in an expression involving $E^T (e^{\xi (z) \langle X \rangle_T})$. We replace this with the expectation of a power option payoff using theorem 6.4.2. Finally, we move the expectation outside the affine function and outside the integral or derivative, yielding an expression of the form $E^T (F_h (S_T))$. This last step may not be valid if the integral is not absolutely convergent, leaving us with an integral of power option prices, instead of a single European option price. However, this is still sufficient to determine the price of the volatility derivative given the complete continuum of call and put prices.

Throughout this section we assume that the price of exponential bond variance contracts are known for all powers $\lambda$. This in turn allows us to treat $F_{\exp} (S_T; \lambda)$ as the payoff function of a European option, whose value equals that of an exponential variance contract. We later give empirical examples using the Hull and White (1990) interest rate model. In practice, as long as bond prices satisfy pull to par, have independent volatility, and $E^T (e^{\lambda Y_T})$ is known and finite, we can apply the results of this chapter.

### 6.5.1 Power Payoffs

The next two propositions show us how to price volatility derivatives with payoffs of the form $h (\langle X \rangle_T) = (\langle X \rangle_T + c)^r$, where $r < 1$ and $c \geq 0$. These are then extended, in propositions 6.5.3 and 6.5.4, to price such derivatives at any time after inception.

**Proposition 6.5.1.** Assume 6.2.1 to 6.2.4. For $0 < r < 1$ and $c \geq 0$,

$$E^T (((X)_T + c)^r) = \frac{r}{\Gamma(1 - r)} \int_0^\infty \frac{1 - e^{-z c} E^T (F_{\exp} (S_T; -z))}{z^{r+1}} dz. \quad (6.5.1)$$
Proof. Following Carr and Lee (2008), for $0 < r < 1$ and $q \geq 0$, we use the following identity:

$$q^r = \frac{r}{\Gamma(1 - r)} \int_0^\infty 1 - e^{-ze^{\langle X \rangle_T + c}} \frac{dz}{z^{r+1}}.$$ 

Setting $q = \langle X \rangle_T + c$ and taking the expectation of both sides yields

$$E^T (((\langle X \rangle_T + c) - c)^r) = E^T \left( \frac{r}{\Gamma(1 - r)} \int_0^\infty 1 - e^{-ze^{\langle X \rangle_T + c}} \frac{dz}{z^{r+1}} \right).$$

$$= \frac{r}{\Gamma(1 - r)} \int_0^\infty 1 - e^{-ze^{\langle X \rangle_T}} \frac{dz}{z^{r+1}}$$

$$= \frac{r}{\Gamma(1 - r)} \int_0^\infty 1 - e^{-ze^{\langle X \rangle_T}} \frac{dz}{z^{r+1}}$$

The exchange of integration and expectation performed above is valid because $\langle X \rangle_T$ is bounded.

\[ \square \]

Proposition 6.5.2. Assume 6.2.1 to 6.2.4. For $r > 0$ and $c > 0$,

$$E^T ((\langle X \rangle_T + c)^{-r}) = \frac{1}{r \Gamma(r)} \int_0^\infty E^T \left( F_{exp}(S_T; -z^{1/r}) \right) e^{-z^{1/r}c} dz. \quad (6.5.2)$$

Proof. Following Carr and Lee (2008), for $r > 0$ and $q > 0$, we use the following identity:

$$q^{-r} = \frac{1}{r \Gamma(r)} \int_0^\infty e^{-z^{1/r}q} dz.$$ 

Setting $q = \langle X \rangle_T + c$ and taking the expectation of both sides yields

$$E^T ((\langle X \rangle_T + c)^{-r}) = E^T \left( \frac{1}{r \Gamma(r)} \int_0^\infty e^{-z^{1/r}(\langle X \rangle_T + c)} dz \right)$$

$$= \frac{1}{r \Gamma(r)} \int_0^\infty E^T \left( e^{-z^{1/r}(\langle X \rangle_T)} \right) e^{-z^{1/r}c} dz$$

$$= \frac{1}{r \Gamma(r)} \int_0^\infty E^T \left( F_{exp}(S_T; -z^{1/r}) \right) e^{-z^{1/r}c} dz.$$ 

The exchange of integration and expectation performed above is valid because $\langle X \rangle_T$ is bounded.

\[ \square \]

Note that the entire effect of stochastic interest rates is contained in the payoff function $F_{exp}(S_T; \lambda)$, defined in theorem 6.4.2. Now to complete the program outlined above, we would like to exchange the order of integration and expectation on the RHS of equations (6.5.1) and (6.5.2). Unfortunately, this is not easy due to the influence of the term $E^T \left( e^{\lambda \langle Y \rangle_T} \right)^{-1}$ in the definition of $F_{exp}(S_T; \lambda)$.

However, if $Q^T (\langle Y \rangle_T \leq c) > 0$, then the RHS of equation (6.5.1) is absolutely convergent, allowing us to exchange the order of integration and expectation. Likewise, if $Q^T (\langle Y \rangle_T < c) > 0$, then
the RHS of equation (6.5.2) is also absolutely convergent. Furthermore, by conditioning on $\mathcal{F}_t$, setting $c = \langle X \rangle_t + d$, and replacing $\langle X \rangle_T$ with $(X)_{t,T} := (X)_{t} - \langle X \rangle_t$, we can use propositions 6.5.1 and 6.5.2 to value power payoffs at any time after inception. This yields the following two propositions, which let us price power payoffs using equations of the form $\mathbb{E}_t^T(h(\langle X \rangle_T)) = \mathbb{E}_t^T(F_h(S_T))$. See appendix 6.A for the proofs.

**Proposition 6.5.3.** Assume 6.2.1 to 6.2.4. Let $d \geq -\langle X \rangle_t$, $0 < r < 1$. If $\mathbb{Q}_t^T(\langle Y \rangle_{t,T} \leq (X)_t + d) > 0$, then

$$
\mathbb{E}_t^T((X)_t + d)^r) = \mathbb{E}_t^T(F_{pow}(S_T; r, d, t)), \text{ where }
F_{pow}(S_T; r, d, t) = \frac{r}{\Gamma(1 - r)} \int_0^\infty 1 - e^{-z(\langle X \rangle_t + d)} \frac{e^{-z(\langle Y \rangle_t)} - 1}{z^{r+1}} \left( \left( \frac{S_T}{F_t} \right)^{p_+} g_1 + \left( \frac{S_T}{F_t} \right)^{p_-} g_2 \right) dz,
$$

and $p_+, p_-, g_1$ and $g_2$ are as in theorem 6.4.2 with $\lambda = -z$.

**Proposition 6.5.4.** Assume 6.2.1 to 6.2.4. Let $d > -\langle X \rangle_t$, $r > 0$. If $\mathbb{Q}_t^T(\langle Y \rangle_{t,T} < (X)_t + d) > 0$, then

$$
\mathbb{E}_t^T((X)_t + d)^{-r}) = \mathbb{E}_t^T(F_{pow}(S_T; -r, d, t)), \text{ where }
F_{pow}(S_T; -r, d, t) = \frac{1}{r \Gamma(r)} \int_0^\infty \mathbb{E}_t^T(e^{-z/\langle Y \rangle_t}) - 1 \left( \left( \frac{S_T}{F_t} \right)^{p_+} g_1 + \left( \frac{S_T}{F_t} \right)^{p_-} g_2 \right) e^{-z^{1/r}(\langle X \rangle_t + d)} dz,
$$

and $p_+, p_-, g_1$ and $g_2$ are as in theorem 6.4.2 with $\lambda = -z^{1/r}$.

For example, consider an interest rate model in which bonds have deterministic volatility, such as the Hull-White model. If the quadratic variation of the log stock price observed to date, $\langle X \rangle_t$, exceeds the predicted quadratic variation of the log bond price from now until expiry, $(X)_t,T$, then the conditions in propositions 6.5.3 and 6.5.4 are satisfied for $d \geq 0$, and we can use the given European options to price power payoffs. Since bond volatility is usually larger further away from expiry, and small compared to stock volatility, we can expect $\langle X \rangle_t$ to exceed $(X)_t,T$ sometime after inception in real world scenarios.

We can also determine the price of positive integer power payoffs by differentiating our expression for $\mathbb{E}_t^T(e^{\lambda\langle X \rangle_T})$ with respect to $\lambda$ and then setting $\lambda$ equal to zero.

**Proposition 6.5.5.** Assume 6.2.1 to 6.2.4. For $n \in \mathbb{Z}^+$,

$$
\mathbb{E}_t^T(\langle X \rangle_t^n) = \mathbb{E}_t^T(F_{pow}(S_T; n)),
$$

where $F_{pow}$ is correlation neutral, and is equal to

$$
F_{pow}(S_T; n) = \frac{\partial^n}{\partial \lambda^n} F_{exp}(S_T; \lambda) \bigg|_{\lambda=0}.
$$
Proof. As \( \langle X \rangle_T \) is bounded, we know that
\[
\mathbb{E}^T (\langle X \rangle_T^n) = \mathbb{E}^T \left( \frac{\partial^n}{\partial \lambda^n} e^{\lambda \langle X \rangle_T} \right)_{\lambda=0} = \frac{\partial^n}{\partial \lambda^n} \mathbb{E}^T (F_{\exp}(S_T; \lambda))_{\lambda=0} = \mathbb{E}^T (F_{\text{pow}}(S_T; n)).
\]

\( F_{\text{pow}} \) is correlation neutral because
\[
\frac{\partial}{\partial y} F_{\text{pow}} (y; n) = \frac{\partial^n}{\partial \lambda^n} \frac{\partial}{\partial y} F_{\exp}(y; \lambda)_{\lambda=0} \Rightarrow \int_0^\infty y \frac{\partial}{\partial y} F_{\exp}(y; \lambda) \phi_{\mu_1, \nu_2} (y) dy \bigg|_{\lambda=0} = 0,
\]
as \( \int_0^\infty y \frac{\partial}{\partial y} F_{\exp}(y; \lambda) \phi_{\mu_1, \nu_2} (y) dy = 0 \) by the proof of theorem 6.4.2.

Setting \( n = 1 \) in the above proposition, we see that the fair strike of a standard variance swap is equal to the undiscounted price of a European option with the following payoff function:
\[
F_{\text{pow}}(S_T; 1) = 2 \left( \frac{S_T}{F_0} - 1 - \log \left( \frac{S_T}{F_0} \right) \right) - \mathbb{E}^T (\langle Y \rangle_T).
\]

6.5.2 Payoffs with Exponentially Decaying Transforms

In general, we can use the inverse Laplace transform to convert our knowledge of \( \mathbb{E}^T (e^{\lambda \langle X \rangle_T}) \) into a value for \( \mathbb{E}^T (h (\langle X \rangle_T)) \). In the case where the Laplace transform of \( h \) decays exponentially we can reverse the order of integration and expectation. As explained previously, this means that we only need to evaluate the price of a single European option, whose payoff takes the form of an integral expression, instead of an infinite set of power options (as is necessary to apply propositions 6.5.1 and 6.5.2).

Proposition 6.5.6. Assume 6.2.1 to 6.2.4. Let \( h : \mathbb{R} \to \mathbb{R} \) be continuous, and \( \alpha \in \mathbb{R} \) such that \( \int_0^\infty e^{-\alpha q} h(q) dq < \infty \). For \( \Re(z) = \alpha \), define \( H(z) := \int_0^\infty e^{-zq} h(q) dq \). Assume that
\[
\left| \mathbb{E}^T (e^{(\alpha+\beta) \langle Y \rangle_T}) \right| \geq k_1 e^{-k_2 |\beta|}
\]
for all \( \beta \in \mathbb{R} \) and some \( k_1, k_2 \geq 0 \). Also assume that \( |H(\alpha + \beta i)| = O(e^{-|\beta|\mu}) \) as \( |\beta| \to \infty \) for some \( \mu > \frac{m_1+m_2}{2} + k_2 \). Then
\[
\mathbb{E}^T (h (\langle X \rangle_T)) = \mathbb{E}^T (F_h(S_T)),
\]
where \( F_h \) is correlation neutral, and equal to
\[
F_h(S_T) = \frac{1}{2\pi i} \int_{\alpha-i\infty}^{\alpha+i\infty} H(z) F_{\exp}(S_T; z) dz.
\]
6.5. Other Variance Contracts

Proof.

\[ E^T(h(\langle X \rangle_T)) = E^T \left( \frac{1}{2\pi i} \int_{\alpha-i\infty}^{\alpha+i\infty} H(z) e^{z\langle X \rangle_T} dz \right), \]
\[ = \frac{1}{2\pi i} \int_{\alpha-i\infty}^{\alpha+i\infty} H(z) E^T(e^{z\langle X \rangle_T}) dz, \]
\[ = \frac{1}{2\pi i} \int_{\alpha-i\infty}^{\alpha+i\infty} H(z) E^T(F_{\exp}(S_T; z)) dz, \]
\[ = E^T \left( \frac{1}{2\pi i} \int_{\alpha-i\infty}^{\alpha+i\infty} H(z) F_{\exp}(S_T; z) dz \right). \]

The first exchange of integration and expectation is justified by our assumption that \( \langle X \rangle_T \) is bounded. Letting \( z = \alpha + \beta i \), the second is justified by

\[ \mathbb{E}^T \left( \left| \left( \frac{S_T}{F_0} \right)^{p_\pm} \right| \right) = \mathbb{E}^T \left( \left| \left( \frac{S_T}{F_0} \right)^{\left( \frac{1}{2} \pm \sqrt{\frac{1}{4} + 2z} \right)} \right| \right), \]
\[ = \mathbb{E}^T \left( \left( \frac{S_T}{F_0} \right)^{\Re \left( \frac{1}{2} \pm \sqrt{\frac{1}{4} + 2(\alpha + \beta i)} \right)} \right), \]
\[ = \mathbb{E}^T \left( \left( \frac{S_T}{F_0} \right)^{\left( \frac{1}{2} \pm \sqrt{\frac{1}{4} + |\beta| + O(|\beta|^{-1/2})} \right)} \right) \text{ as } \beta \to \pm \infty, \]
\[ = \mathbb{E}^T \left( e^{\left| \beta \right| + O(1)} \langle (X)_T + (Y)_T \rangle \right), \]
\[ = \mathcal{O} \left( e^{\left| \beta \right| (m_1 + m_2)} \right), \]

and

\[ \left| H(z) E^T \left( e^{z\langle Y \rangle_T} \right)^{-1} g_k \right| = \mathcal{O} \left( e^{-|\beta|(\mu - k_2)} \right) \text{ as } \beta \to \pm \infty, \]

where \( g_k = g_1 \) or \( g_2 \), and \( \mu > \frac{m_1 + m_2}{2} + k_2 \). Therefore the integral is absolutely convergent. Also, correlation neutrality follows easily from that of \( F_{\exp}(S_T; z) \).

Although many payoff functions of interest, such as volatility swaps and volatility puts, do not have Laplace transforms which decay exponentially, Friz and Gatheral (2005) have explained how to construct smoothed versions of such payoff functions that do in fact satisfy this condition. However, their smoothing procedure may lead to numerical difficulties as the approximation becomes more accurate. We refer the reader to their paper for details.

6.5.3 Other Payoff Functions

Carr and Lee (2008) showed that the price of a general option on variance, which pays \( h(\langle X \rangle_T) \) at time \( T \), is determined by knowledge of \( \mathbb{E}^T(e^{\langle X \rangle_T}) \) for a wide range of payoff functions not covered
above, including those with integrable Laplace transforms and those that are continuous on $[0, \infty]$. These results rely only on the boundedness of $\langle X \rangle_T$, and not on the particular form of the expression available for $E^T(e^{\lambda \langle X \rangle_T})$. Therefore we can adapt them to stochastic interest rates by using our formula, found in theorem 6.4.2, for the price of an exponential stock variance contract. However, their application requires evaluating an infinite number of power option prices, instead of just one specific non-path-dependent option.

### 6.6 Unbounded Quadratic Variation

Under assumptions 6.2.1 to 6.2.4, propositions 6.5.3 to 6.5.6 tell us how to price volatility derivatives using equations of the form $E^T(h(\langle X \rangle_T)) = E^T(F_h(S_T))$. By extending the approach of Carr and Lee (2008), the theorem below shows that these propositions are still valid without assumption 6.2.4, as long as $h$ and $F_h$ satisfy certain additional conditions.

**Theorem 6.6.1.** Suppose that there exists measurable functions $h$ and $F_h$ such that

$$E^T(h(\langle X \rangle_T)) = E^T(F_h(S_T)),$$

for all models satisfying assumptions 6.2.1 to 6.2.4. Assume that $F_h$ can be written as $F_h = F_+ - F_-$, where $F_\pm$ is convex and $E^T(F_\pm(S_T)) < \infty$. Further assume that either

(i) $h$ is bounded, or  
(ii) $h$ non-negative and increasing.

Then equation (6.6.1) also holds for all models that satisfy $E^T(\langle X \rangle_T) < \infty$ and assumptions 6.2.1 to 6.2.3.

**Proof.** Let $m \in \mathbb{R}^+$. Given a model satisfying assumptions 6.2.1 to 6.2.3, define $\tau_\eta := \inf\{s : \langle X \rangle_s \geq m\}$ and $\tau_\psi := \inf\{s : \langle Y \rangle_s \geq m\}$. Consider an altered model defined as follows:

$$\eta_t^* := \eta_t \mathbb{1}_{t<\tau_\eta}, \quad \psi_t^* := \psi_t \mathbb{1}_{t<\tau_\psi},$$

$$r_t^* := r_t \mathbb{1}_{t<\tau_\psi} - \frac{1}{T - \tau_\psi} \log(P_{\tau_\psi}) \mathbb{1}_{t \geq \tau_\psi},$$

$$dS_t^* = r_t^* S_t^* + \eta_t^* S_t^* dW_t^{S,T},$$

$$dP_t^* = \left(r_t^* + (\psi_t^*)^2\right) P_t^* dt + \psi_t^* P_t^* dW_t^{P,T},$$

$$X_t^* := \log(S_t^*), \quad Y_t^* := \log(Y_t^*).$$

This altered model satisfies assumptions 6.2.1 to 6.2.4. Thus $E^T(h(\langle X^* \rangle_T)) = E^T(F_h(S_T^*))$ for all $m \in \mathbb{R}^+$. Taking the limit as $m \to \infty$, we see that $\langle X^* \rangle_T \xrightarrow{a.s.} \langle X \rangle_T$. Therefore, by either the dominated convergence theorem in case (i), or the monotone convergence theorem in case (ii), we know that

$$\lim_{m \to \infty} E^T(h(\langle X^* \rangle_T)) = E^T(h(\langle X \rangle_T)).$$
To complete the proof, we must show that \(\mathbb{E}^T(F_h(S_T)) \to \mathbb{E}^T(F_h(S_T))\) as \(m \to \infty\). This follows from the zero-interest-rates special case presented in Carr and Lee (2008), by replacing \(S_0\) with \(F_0\), and conditioning on \((\langle X \rangle_T, \langle Y \rangle_T)\) instead of \(\langle X \rangle_T\). We refer the reader to the last page of appendix A in their paper for details.

Carr and Lee (2008) have observed that the condition on \(F_h\) above is true if \(F_h\) can be replicated with calls and puts using a long position of finite price and a short position of finite price, and is therefore quite mild.

### 6.7 Mixture of Normals Method

Under independent interest rates, the specific choice of the process for the bond price will have no effect on the distribution of variance, \(\langle X \rangle_T\), in the T-forward measure. Therefore any two models with the same process for stock volatility in that measure will give the same prices for volatility derivatives, regardless of the interest rate process, as long as the T-expiry zero coupon bond prices are the same.

Furthermore, the methods discussed above require knowledge of the complete continuum of vanilla option prices. However, only a finite set of prices are observable in the market, and these may be noisy due to a lack of liquidity and the bid-ask spread. Thus, a practical approach to pricing volatility derivatives is to fit a curve to the market smile under a given interest rate model, set interest rate stochasticity to zero while holding bond prices constant, and then use the methods of Carr and Lee (2008) or Friz and Gatheral (2005).

To illustrate this approach we will use a mixture of normals to approximate the density of the log stock price at expiry in the T-forward measure, assuming that the integrated short rate, \(\int_0^T r_u \, du\), also has a normal mixture distribution. We then determine what distribution the final stock price would have if interest rates were made deterministic, and use this to price volatility derivatives.

The mixture of normals method, as presented by Alexander (2001), directly models the risk-neutral density of the final log stock price as a weighted sum of normal densities. The mean of each density is set to the same value, based on the risk free interest rate, to ensure that the expected value of the discounted stock price is equal to the initial price. Then, the weights and variances are found by minimizing the sum of the squared deviations of the model call option prices from the observed market prices, subject to the condition that the weights are positive and sum to one. Rebonato and Cardoso (2004) extended this approach by allowing the means to take differing values, as this is necessary to fit a skew in implied volatilities. They also showed how the minimization can be performed in an unconstrained manner by writing the weights as the squares of the coordinates of a point on the unit hyper sphere.

To make our approach explicit, we first write the final stock price, \(S_T\), as the product of two
independent parts: the bank account, \( \beta_T := e^{\int_0^T r_u \, du} \), and the exponential martingale,
\[
\tilde{S}_T := S_0 e^{-\frac{1}{2} \int_0^T \eta_u^2 \, du + \int_0^T \eta_u \, dW^u_S}.
\]
Here, \( \eta_u \) is an unknown stochastic volatility process, independent of the bank account, but correlated with the driving Brownian motion, \( W^S_T \). Thus, we have that \( S_T = \beta_T \tilde{S}_T \), and we approximate the densities of \( \beta_T \) and \( \tilde{S}_T \) with the log-normal mixture distributions
\[
\phi_{\beta_T}(x) = \sum_{j=1}^{m_1} a_j \gamma_j(x), \text{ where } \gamma_j = \mathcal{LN} \left( -\log(P_0) + \left( \mu_j - \frac{1}{2} \sigma_j^2 \right) T, \sigma_j^2 T \right), \text{ and } \\
\phi_{\tilde{S}_T}(x) = \sum_{k=1}^{m_2} b_k \zeta_k(x), \text{ where } \zeta_k = \mathcal{LN} \left( \log(S_0) + \left( \mu_k - \frac{1}{2} \sigma_k^2 \right) T, \sigma_k^2 T \right).
\]
Under these two approximations, \( S_T \) also has a log-normal mixture distribution. Therefore, given the interest rate model parameters \( (a_j, \bar{\mu}_j \text{ and } \bar{\sigma}_j^2) \), we can determine the parameters \( (b_k, \mu_k \text{ and } \sigma_k^2) \) for the distribution of \( \tilde{S}_T \) that best fits the observed vanilla option prices by using a least squares optimization as described by Rebonato and Cardoso (2004). This fit must be performed with the following non-linear restriction on \( \mu_k \):
\[
\mathbb{E}^T \left( \tilde{S}_T \right) = \sum_{k=1}^{m_2} b_k e^{\log(S_0) + \mu_k T} = S_0. \\
\implies \sum_{k=1}^{m_2} b_k e^{\mu_k T} = 1
\]
However, because the terms \( b_k e^{\mu_k T} \) are each positive and sum to one, we can use the same transformation to deal with the restriction on the means, \( \mu_k \), that Rebonato and Cardoso (2004) introduced to deal with the restriction on the weights, \( b_k \). In other words, we set each one to the square of a coordinate of the unit hyper sphere, as follows:
\[
\mu_k = \frac{1}{T} \log \left( \frac{\alpha_k^2}{b_k} \right), \text{ where } \\
\alpha_k = \cos(\theta_k) \prod_{j=1}^{k-1} \sin(\theta_j), k = 1, 2, ..., m_2 - 1, \text{ and } \\
\alpha_{m_2} = \prod_{j=1}^{m_2-1} \sin(\theta_j), \theta_j \in \left( 0, \frac{\pi}{2} \right).
\]
Note that the Hull and White (1990) interest rate model is a special case of the previous setup, with \( m_1 = 1, a_1 = 1, \bar{\mu}_1 = 0 \) and \( \bar{\sigma}_j^2 \) a deterministic function of the Hull-White parameters. We will use this model later in the chapter to give some empirical examples of the effect of stochastic interest rates on a range of volatility derivatives.
Now we know that the distribution of $\tilde{S}_T$ determines the price of volatility derivatives, regardless of the distribution of the bank account, $\beta_T$, as long as the price of the T-expiry zero coupon bond remains unchanged. We exploit this fact by considering an alternative model, $S^*_T$, for the stock price, in which the interest rate has been made deterministic by replacing $\beta_T$ with $P_{0}^{-1}$. Thus we write $S^*_T = P_{0}^{-1}\tilde{S}_T$, and approximate its density using

$$
\phi_{S^*_T}(x) = \sum_{k=1}^{m_2} b_k \xi_k(x), \text{ where } \xi_k = \mathcal{LN}\left(\log (F_0) + \left(\mu_k - \frac{1}{2} \sigma_k^2\right) T, \sigma_k^2 T\right).
$$

We will now use this distribution to price volatility derivatives using methods which assume that interest rates are deterministic, with the knowledge that the price will be the same as that implied by the full model, which includes independent stochastic interest rates. For example, by proposition 6.5.5 with $\langle Y\rangle_T = 0$, the correlation neutral estimate of the fair strike of a variance swap is

$$
\mathbb{E}^T(\langle X\rangle_T) \approx \sum_{k=1}^{m_2} b_k (\sigma_k^2 - 2\mu_k).
$$

Likewise, by proposition 6.5.1 with $\langle Y\rangle_T = 0$, the correlation neutral estimate of the price of a volatility swap is

$$
\mathbb{E}^T\left(\sqrt{\langle X\rangle_T + c}\right) \approx \frac{1}{2\sqrt{\pi}} \int_{0}^{\infty} \frac{1 - e^{-2c\mathbb{E}^T(F_{\exp}(S^*_T;-z))}}{z^{3/2}} dz, \text{ where}
$$

$$
\mathbb{E}^T(F_{\exp}(S^*_T;-z)) = \sum_{k=1}^{m_1} b_k e^{-z\sigma_k^2 T}\left(g_1 e^{\left(\frac{1}{2} + \sqrt{\frac{1}{4} - 2z}\right)\mu_k T} + g_2 e^{\left(\frac{1}{2} - \sqrt{\frac{1}{4} - 2z}\right)\mu_k T}\right).
$$

Here $g_1$ and $g_2$ are as in theorem 6.4.2, with $\lambda = -z$. In the case that $\mu_k$ is equal to zero for all $k$, these formulas reduce to the weighted average of the pay-off function evaluated at the discrete set of variances, $\sigma_k^2$. This is because the log-normal mixture distribution assigned to $S^*_T$ can be generated by any model in which $\langle X\rangle_T$ has a discrete distribution with point mass $b_k$ at value $\sigma_k^2$ for $k = 1, 2, ..., m_1$.

As explained by Rebonato and Cardoso (2004), fitting a skew in the implied volatility smile requires non-zero values of $\mu_k$. However, the zero correlation assumption regarding the stock price volatility process and driving Brownian motion yields a form of put-call symmetry which results in the equality

$$
\mathbb{E}^T\left(\frac{S^*_T}{F_0}\right)^{\frac{1}{2} + \sqrt{\frac{1}{4} + 2\lambda}} = \mathbb{E}^T\left(\frac{S^*_T}{F_0}\right)^{\frac{1}{2} - \sqrt{\frac{1}{4} + 2\lambda}}
$$

for all $\lambda \in \mathbb{C}$. This equality is not satisfied when some $\mu_k$ are not equal to zero, implying that the zero correlation assumption is false. Thus, the use of the first-order correlation immune estimate is important when pricing volatility derivatives based on fits to vanilla option prices that display a skew in implied volatility.
6.7.1 Comparison to Existing Fitting Procedures

Friz and Gatheral (2005) have explained how the distribution of the square root of quadratic variation, \( \sqrt{\langle X \rangle_T} \), can be fit directly to observed call and put prices. They began by approximating it using the following log-normal distribution, whose parameters are specifically chosen to match the fair strikes of a volatility swap and a variance swap.

\[
\log \left( \sqrt{\langle X \rangle_T} \right) \sim \mathcal{N}(m, s^2), \quad \log (\langle X \rangle_T) \sim \mathcal{N}(2m, 4s^2),
\]

\[
m = \frac{1}{2} \log \left( \frac{E_T \left( \sqrt{\langle X \rangle_T} \right)^4}{E_T \left( \langle X \rangle_T \right)^2} \right), \quad s^2 = \frac{1}{2} \log \left( \frac{E_T \left( \sqrt{\langle X \rangle_T} \right)^2}{E_T \left( \langle X \rangle_T \right)^2} \right).
\]

In the case of deterministic interest rates, the terms \( E_T \left( \langle X \rangle_T \right) \) and \( E_T \left( \sqrt{\langle X \rangle_T} \right) \), which are the fair strikes of a variance swap and a volatility swap respectively, can be estimated from the observed market prices of vanilla options using the techniques of Carr and Lee (2008). In the case of stochastic rates, we find them using equations (6.7.1) and (6.7.2) respectively.

Friz and Gatheral then discretized the distribution of \( \sqrt{\langle X \rangle_T} \) using a finite set of log-volatilities, \( z_i \), each occurring with probability \( q_i \), \( i = 1, 2, ..., n_{\text{vol}} \), where \( z_i \) and \( q_i \) are chosen to accurately represent a \( \mathcal{N}(m, s^2) \) distribution. Finally, they assigned a posterior probability, \( p_i \), to each log-volatility level, \( z_i \), by minimizing the following objective function:

\[
\mathcal{O}(p) = \sum_{j=1}^{n_{\text{strikes}}} \left( \sum_{i=1}^{n_{\text{vol}}} p_i c_{BS} \left( K_j, e^{2z_i} \right) \right)^2 - c(K_j) + \beta d(p, q),
\]

where \( \sum_{i=1}^{n_{\text{vol}}} p_i c_{BS} \left( K_j, e^{2z_i} \right) \) is the model price of a call option with strike \( K_j \), \( c(K_j) \) is the market price, \( \beta \) is a non-negative constant, and \( d(p, q) \) is a given measure of the distance between probability vectors \( p \) and \( q \), such as the relative entropy distance. This objective function measures the sum of the squared difference between the model and market prices of call options, plus a penalty proportional to the distance between the prior and posterior probability vectors. Note that the expression given for the model price of a call option is an application of the well-known result due to Hull and White (1987), which says that the price of a non-path-dependent derivative in an uncorrelated stochastic volatility model can be found by conditioning on the variance path and then integrating against the density of total variance.

In the special case that \( \beta \) is set to zero, so that no weight is given to the log-normal prior distribution, the Friz and Gatheral approach gives the same results as our mixture model approach under the assumption of deterministic interest rates, fixed volatility levels, \( \sigma_k \sqrt{T} = e^{\sigma_k} \), and zero skew (i.e. \( \mu_k = 0 \) for all \( k \)).
6.8 Empirical Results

In this section we estimate the impact of stochastic interest rates on the value of variance swaps and other volatility derivatives. Starting with equation (6.3.2), we substitute \( \rho = \frac{1}{2} - \sqrt{\frac{1}{4} + 2\lambda} \) and take the derivative with respect to \( \lambda \). Setting \( \lambda = 0 \) and rearranging yields

\[
\mathbb{E}_T \left( \langle X \rangle_T \right) = -2 \mathbb{E}_T \left( \log \left( \frac{S_T}{F_0} \right) \right) - \mathbb{E}_T \left( \langle Y \rangle_T - 2 \int_0^T \rho \psi_u \eta_u du \right). \tag{6.8.1}
\]

The first term on the right hand side is the undiscounted price of a European option, and is therefore fully determined by a given discount curve and call/put price continuum. Thus, the effect on variance swaps of introducing stochastic rates to the model, while holding bond and vanilla option prices constant, is due entirely to the second term, which is zero when interest rates are deterministic.

Hörfelt and Torné (2010) derived equation (6.8.1) by analysing, under stochastic interest rates, the relationship between a log contract and the payoff of a variance swap. This analysis, which only applies to variance swaps, and not other types of volatility derivatives, was followed by an empirical example, based on the S&P index as at 8th December 2008, using an equity and interest rate hybrid model found in Overhaus et al. (2007). In the case of independent interest rates, i.e. \( \rho = 0 \), they found that interest rate stochasticity has almost no effect on the fair strike of variance swaps with a term of 1 or 2 years, and causes only a 0.20% and 0.35% relative reduction for 3 and 4 year swaps respectively. Note that we have inferred these relative changes based on the fair strike quoted in variance points, \( \mathbb{E}_T \left( \langle X \rangle_T \right) \), not in volatility points, \( \sqrt{\mathbb{E}_T \left( \langle X \rangle_T \right)} \).

However, the magnitude of the effect of interest rate stochasticity may be significantly larger for longer term contracts. Using market data from Thomson Reuters Datastream, we applied the approach detailed in section 6.7 to calculate the fair strikes of volatility swaps and variance swaps, and also the prices of call or puts on volatility or variance. We valued 9.15 year contracts, starting on 21 October 2014 and written on the Euro Stoxx 50 index, by fitting a Hull-White model for interest rates to Euro cap volatility quotes, and then fitting a normal mixture model with 4 basis functions to the market prices of call and put options expiring on 15 December 2023.

Variance swaps and volatility swaps were valued using equations (6.7.1) and (6.7.2) respectively, while calls and put were valued using the log-normal approximation described by equation (6.7.3). Looking at table 6.1, we see that the introduction of stochastic interest rates causes a 1.3% and 2.0% drop in the fair strikes of volatility and variance swap respectively. Likewise, we see significant deviations in the prices of at the money variance and volatility options, ranging from -2.6% for variance calls to +4.4% for volatility puts. The signs of these deviations are explained by the fact that the fair strikes of the associated swaps under stochastic rates are lower than the fair strikes under deterministic rates. Therefore, holding the strike constant, a put which is at the money under stochastic rates will be out of the money under deterministic rates.

In the \( \rho = 0 \) case, the most important consideration is the relative sizes of expected stock vari-
Table 6.1: Fair strikes and prices of volatility derivatives valued at 21 October 2014 and expiring on 15 December 2023

<table>
<thead>
<tr>
<th>Type</th>
<th>Deterministic rates</th>
<th>Stochastic rates</th>
<th>% Adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility swap fair strike</td>
<td>0.1960</td>
<td>0.1935</td>
<td>-1.3%</td>
</tr>
<tr>
<td>Variance swap fair strike</td>
<td>0.0477</td>
<td>0.0467</td>
<td>-2.0%</td>
</tr>
<tr>
<td>ATM volatility call price</td>
<td>0.1047</td>
<td>0.1018</td>
<td>-2.8%</td>
</tr>
<tr>
<td>ATM volatility put price</td>
<td>0.0975</td>
<td>0.1018</td>
<td>4.4%</td>
</tr>
<tr>
<td>ATM variance call price</td>
<td>0.1485</td>
<td>0.1446</td>
<td>-2.6%</td>
</tr>
<tr>
<td>ATM variance put price</td>
<td>0.1402</td>
<td>0.1446</td>
<td>3.2%</td>
</tr>
</tbody>
</table>

1. Volatility swap fair strike quoted in annualized volatility points.
2. Variance swap fair strike quoted in annualized variance points.
3. The ATM strike is set as the fair strike of the associated swap under stochastic rates.

...  

6.9 Conclusion

Previous work on the pricing of volatility derivatives has generally assumed that interest rates are deterministic. Some of the more recent papers in the area have considered the effect of stochastic interest rates, but have been either model dependent or only applicable to variance swaps. Building on the innovative model-free work of Carr and Lee (2008), we have shown that, under a class of stochastic interest rate models, the prices of a wide range of volatility derivatives are determined given the complete continuum of vanilla option prices and a calibrated interest rate process.

Our key assumptions are that the bond and stock price processes are continuous, instantaneously uncorrelated with each other, and have volatility processes that are independent of their driving Brownian motions. Unlike previous work, our results are model free, apply to general volatility derivatives, and account for non-deterministic interest rates. By model free, we mean that we do not require a specific process for the stock price or its volatility, but we do make some strong assumption regarding the nature of the interest rate process and the independence of various factors.

We have also given pricing algorithms, which are first-order immune to the presence of correlation between the stock’s volatility process and driving Brownian motion. This includes an extension of the method of Rebonato and Cardoso (2004) to allow for independent stochastic interest rates when estimating the density of the final stock price based on market vanilla option prices. Then, by considering an alternative model in which interest rates are deterministic but volatility derivatives prices are unchanged, we are able to apply existing results which were previously only accurate under deterministic rates.

Finally, we have examined the empirical impact of interest rate stochasticity on variance swaps,
volatility swaps, and calls or puts on variance or volatility. In doing so, we have found that there is indeed a significant effect on long-dated contracts, which should not be ignored.

Appendix 6.A  Proofs

Proof of proposition 6.5.3. Starting with proposition 6.5.1, we condition on \( F_t \), set \( c = \langle X \rangle_t + d \), and replace \( \langle X \rangle_T \) with \( \langle X \rangle_T := \langle X \rangle_T - \langle X \rangle_t \) to get

\[
\mathbb{E}_t^T ((\langle X \rangle_T + d)^r) = \frac{r}{\Gamma(1-r)} \int_0^\infty 1 - e^{-z(\langle X \rangle_t + d)} \mathbb{E}_t^T (F_{exp}(S_T; -z, t)) \, dz,
\]

where

\[
F_{exp}(S_T; \lambda, t) := \mathbb{E}_t^T \left( e^{\lambda(Y)_{t,T}} \right)^{-1} \left( \frac{S_T}{F_t} \right)^{p_1} \mathbb{E}_t^T \left( \frac{S_T}{F_t} \right) \mathbb{E}_t^T \left( \frac{S_T}{F_t} \right)^{p_2} g_1 \mathbb{E}_t^T \left( \frac{S_T}{F_t} \right)^{g_2},
\]

and \( p_+, p_-, g_1 \) and \( g_2 \) are functions of \( \lambda \) as in theorem 6.4.2. We set \( \lambda = -z \), and define \( p_1 := p_+ \), \( p_2 := p_- \), and \( B(z) := e^{-z(\langle X \rangle_t + d)} \mathbb{E}_t^T (e^{-z(Y)_{t,T}})^{-1} \) to get

\[
\mathbb{E}_t^T ((\langle X \rangle_T + d)^r) = \frac{r}{\Gamma(1-r)} \int_0^\infty \mathbb{E}_t^T \left( \sum_{k=1}^2 g_k \frac{1 - B(z) \left( \frac{S_T}{F_t} \right)^{p_k}}{z^{r+1}} \right) \, dz,
\]

as \( g_1 + g_2 = 1 \). In order to complete the proof we will reverse the order of integration and expectation in the above expression by showing that it is absolutely convergent. To this end, let \( p_k := \frac{1}{2} - (-1)^k \sqrt[3]{1 - 2z} \), and

\[
A_k := \mathbb{E}_t^T \left( \left[ 1 - B(z) \left( \frac{S_T}{F_t} \right)^{p_k} \right]^2 \right).
\]

Defining \( q := \mathbb{Q}_t^T (\langle Y \rangle_{t,T} \leq \langle X \rangle_t + d) \), the condition \( q \mathbb{Q}_t^T (\langle Y \rangle_{t,T} \leq \langle X \rangle_t + d) > 0 \) implies that

\[
\mathbb{E}_t^T \left( e^{-z(Y)_{t,T}} \right) \geq q e^{-z(\langle X \rangle_t + d)} \forall z \geq 0.
\]

That is \( B(z) \leq \frac{1}{q} \forall z \geq 0 \). Therefore, for large \( z \), we know that

\[
A_k^2 \leq \mathbb{E}_t^T \left( 1 + B(z) \left( \frac{S_T}{F_t} \right)^{R(p_k)} \right) = 1 + B(z) \mathbb{E}_t^T \left( \left( \frac{S_T}{F_t} \right)^{\frac{1}{2}} \right) = \mathcal{O}(1) \text{ as } z \to \infty.
\]

Furthermore, \( |g_k| = \frac{1}{2} + (-1)^k \frac{1}{2\sqrt{1 - 8z}} = \mathcal{O}(1) \text{ as } z \to \infty \). Therefore

\[
\mathbb{E}_t^T \left( \frac{1 - B(z) \left( \frac{S_T}{F_t} \right)^{p_k} \frac{1}{z^{r+1}}} \right) = \mathcal{O} \left( z^{-r-1} \right) \text{ as } z \to \infty.
\]
Now, for small $z$, we know that $p_k$ is real, and thus
\[
A_k \leq E_t^T \left( \left| 1 - B(z) \left( \frac{S_T}{F_t} \right)^{p_k} \right|^2 \right) \\
= E_t^T \left( 1 - 2B(z) \left( \frac{S_T}{F_t} \right)^{p_k} + B^2(z) \left( \frac{S_T}{F_t} \right)^{2p_k} \right).
\]

Defining $\theta_k := \frac{1-8z-(1)^k\sqrt{1-8z}}{2}$, we can write $2p_k$ as
\[
2p_k = 2 \left( \frac{1}{2} - (-1)^k \sqrt{\frac{1}{4} - 2z} \right) = \frac{1}{2} - (-1)^k \sqrt{\frac{1}{4} + 2\theta_k}.
\]

Therefore, using the fact that $E_t^T \left( \left( \frac{S_T}{F_t} \right)^{\frac{1}{2} \pm \sqrt{\frac{1}{4} + 2\lambda}} \right) = E_t^T \left( e^{\lambda(X)_{t,T} + (Y)_{t,T}} \right)$, we have
\[
A_k \leq 1 - 2B(z)E_t^T \left( e^{-z(X)_{t,T} + (Y)_{t,T}} \right) + B^2(z)E_t^T \left( e^{\theta_k(X)_{t,T} + (Y)_{t,T}} \right).
\]

We see that $A_1 = O(1)$ as $z \to 0^+$. Furthermore, we know that the moment generating function $M(z) := E_t^T \left( e^{zI(X)_{t,T} + (Y)_{t,T}} \right)$, $z \in \mathbb{R}$, is analytic because $(X)_{t,T} + (Y)_{t,T}$ is bounded. Likewise, $B(z) := e^{-z(X)_{t,T} + (Y)_{t,T}}$ is analytic because $(Y)_{t,T}$ is bounded. Therefore we can expand about $z = 0$ to get
\[
A_2 = 1 - 2 \left( 1 - zM'(0) + zB'(0) + O(z^2) \right) + 1 - 2zM'(0) + 2zB'(0) + O(z^2)
= O(z^2) \text{ as } z \to 0^+.
\]

Combining this with the fact that $|g_1| = \left| \frac{1}{2} - \frac{1}{2\sqrt{1-zz}} \right| = O(z)$ and $|g_2| = \left| \frac{1}{2} + \frac{1}{2\sqrt{1-zz}} \right| = O(1)$ as $z \to 0^+$, we have
\[
E_t^T \left| \frac{1}{z^{r+1}} \left( \frac{S_T}{F_t} \right)^{p_k} \right| = \frac{|g_k| A_k^2}{z^{r+1}} = O \left( z^{-r} \right) \text{ as } z \to 0^+.
\]  

(6.A.3)

Given that $0 < r < 1$, the bounds in equations (6.A.2) and (6.A.3) show that the integral in equation (6.A.1) is indeed absolutely convergent. Therefore we can reverse the order of integration and expectation as required.

**Proof of proposition 6.5.4.** Starting with proposition 6.5.2, we condition on $F_t$, set $c = \langle X \rangle_t + d$, and replace $\langle X \rangle_T$ with $\langle X \rangle_{t,T} := \langle X \rangle_T - \langle X \rangle_t$ to get
\[
E_t^T \left( (\langle X \rangle_T + d)^{-r} \right) = \frac{1}{r^\Gamma(r)} \int_0^\infty E_t^T \left( F_{exp \left( \langle S_T \rangle_t - z^{1/r}, t \right)} \right) e^{-z^{1/r}(\langle X \rangle_t + d)} dz.
\]
where
\[ F_{\exp}(ST; \lambda, t) := \mathbb{E}_t^T \left( e^{\lambda (Y)_{t,T}} \right)^{-1} \left( \left( \frac{ST}{F_t} \right)^{p_+} g_1 + \left( \frac{ST}{F_t} \right)^{p_-} g_2 \right) , \]
and \( p_+, p_-, g_1 \) and \( g_2 \) are functions of \( \lambda \) as in theorem 6.4.2. We set \( \lambda = -z_1/r \), and define \( p_1 := p_+ \), \( p_2 := p_- \), and \( D(z) := e^{-z^{1/r}(X)_{t+d}} \mathbb{E}_t^T \left( e^{-z^{1/r}(Y)_{t,T}} \right)^{-1} \) to get
\[
\mathbb{E}_t^T ((X)_{T+d})^{-r} = \frac{1}{r \Gamma (r)} \int_0^\infty D(z) \mathbb{E}_t^T \left( \left( \frac{ST}{F_t} \right)^{p_1} \left( \frac{ST}{F_t} \right)^{p_2} \right) dz
\]
(6.A.4)

For large \( z \), we know that
\[
\mathbb{E}_t^T \left( \frac{ST}{F_t} \right)^{p_k} = \mathbb{E}_t^T \left( \frac{ST}{F_t} e^{R(p_k)} \right) = \mathbb{E}_t^T \left( \frac{ST}{F_t} \right)^{\frac{3}{2}} = O(1) \text{ as } z \to \infty.
\]
Furthermore, \( |g_k| = \left| \frac{1}{2} + (-1)^k \frac{1}{2\sqrt{1-8z}} \right| = O(1) \text{ as } z \to \infty. \]
Now, defining \( q(\delta) := \mathbb{Q}_t^T ((Y)_{t,T} < (X)_t + d - \delta) \), the condition \( \mathbb{Q}_t^T ((Y)_{t,T} < (X)_t + d) > 0 \) implies that there exists some \( \delta > 0 \) such that \( q(\delta) > 0 \) and
\[
\mathbb{E}_t^T \left( e^{-z^{1/r}(Y)_{t,T}} \right) \geq q(\delta) e^{-z^{1/r}(X)_{t+d-\delta}}, \forall z \geq 0.
\]
That is \( D(z) \leq q(\delta)^{-1} e^{-z^{1/r}\delta}, \forall z \geq 0. \) Therefore the integral in equation (6.A.4) is absolutely convergent, and we can exchange the order of integration and expectation as required. \( \square \)
Chapter 7

Summary and Conclusion

Since the introduction of the Black-Scholes option pricing model, much work has gone into relaxing its assumptions and overcoming its limitations. Within the class of continuous models in frictionless markets, the primary focus has been on local volatility, stochastic volatility and stochastic interest rates. Individually, each of these proposed extensions deals with an important underlying issue and has received much attention in the literature. The need to accurately price increasingly long-term and complex exotic derivatives has further led to development of more advanced models that combine these extensions together.

However, these more recent models have typically relied on approximations and potentially slow numerical techniques. Although this may be fine when dealing with short-term derivatives, the problems are compounded as the term increases. Thus, there is a need for new modelling approaches which are exact, and enable rapid calibration and simulation. To this end, in chapter 3, we developed a technique for specifying parametric local volatility models under stochastic interest rates. Under this approach, the stock price, adjusted for interest rates and dividends, is written as function of a normal random variable. This stock pricing function is given by the expected value, conditional on the information up to time \( t \), of a chosen payoff function. The flexibility of this method lies in the fact that any increasing payoff function with an analytically known expectation can be used to create the stock pricing function. Thus, quite complex stock pricing functions can be built up out of linear combinations of more simple functions.

A key advantage of these models is that the joint density of the stock price and the short-rate is known analytically. Furthermore, the prices of vanilla options can be computed rapidly by evaluating a one-dimensional integral, which allows the models to be calibrated efficiently. Computation of the local volatility function implied by these models requires evaluating the inverse of the stock...
pricing function. Although this can be done analytically in some simple cases, more complex cases have to be done using a numerical algorithm, such as the Newton-Raphson method. Nevertheless, this does not present a problem because the stock pricing function is typically smooth and monotonic, and the inverse can be cached if necessary. Furthermore, the process for the stock price and the short-rate can be exactly simulated over arbitrary time steps because their joint distribution is a known transformation of the bivariate normal distribution. Thus, there is no need to use a short time-step SDE discretization scheme, which is potentially slow and inaccurate for long-expiries.

Nonetheless, these models have only one time-dependent parameter, namely the volatility of the driving Gaussian process. For this reason, it may be difficult to fit the implied volatility surface for multiple expiries. In chapter 4 we address this issue by showing how to construct multivariate mixture models under stochastic interest rates. Under this approach the joint density of all the state variables, in the forward measure, is equated to a linear combination of component joint densities. Thus, the prices of vanilla option, and in fact any option whose payoff only depends on the values of the state-variables at a single fixed point in time, are easily computed using the linear combination of the prices in each component model.

The drift and instantaneous covariance matrix necessary to produce the desired joint density is equal to a weighted average of the component drifts and instantaneous covariance matrices. Note that the drift of the stock price in each component model is set equal to the short-rate minus the dividend rate all multiplied by the stock price (i.e. \((r_t - y_t)S_t\)), so that the drift is correct in the final model. Furthermore, the same short-rate process is used in each component model, so that the interest rate part of the model is also correct in the final model. The weights in the expressions for the drift and instantaneous covariance matrix depend on the joint densities of the state variables in each component model. Thus the parametric models introduced in chapter 3 are good candidates to use as component models, because this joint density is available in a closed form.

Via an example calibration to FTSE 100 index option data, we saw that a mixture model can produce highly accurate fits to multiple expiries in a real world scenarios. However, even mixing shifted exponential models may not produce enough skew to match the implied volatility curve, which means that we may need to use component models compatible with a higher degree of skew, such as the modified exponential model. The empirical example was also used to examine the effects of stochastic interest rates on long-dated path-dependent options, specifically up-and-out call options. We saw that the level of instantaneous correlation between the short-rate and the stock price can have a large effect on the value of such options, even though they have no explicit hybrid features.

However, local volatility on its own does not produce realistic volatility dynamics, and may not be suitable when pricing exotic contracts whose value depends on the random nature of volatility itself. For this reason, in chapter 5, we extended our mixture model approach to include stochastic volatility. This allows us to calibrate LSVSR models without having to resort to any approximations, Markovian projections, asymptotic expansions, Monte-Carlo simulations or the numerical solution of three-dimensional second-order PDEs, all of which may be too slow or inaccurate to apply to
long-term models.

After the mixture model is calibrated, the joint density of the stock price, short-rate, and volatility process in each component model is cached by using the multidimensional fractional FFT to invert the joint characteristic function. Unlike a standard FFT, this algorithm provides an efficient way to perform the inversion without wasting any calculations in the extreme tails of the input characteristic function or output density. Once the caching process is complete, the component joint densities are used to compute the drift and diffusion coefficients necessary to simulate the underlying SDE.

Our proposed procedure is demonstrated using a mixture of Schöbel-Zhu-Hull-White models. We have extended this component model to allow for piecewise-constant time-dependent parameters, as this helps when calibrating to multiple expiries. Unlike the LVSR models of chapters 3 and 4, the key parameters determining the implied volatility smile at each expiry are time-dependent. This means that the model can be calibrated step-by-step, one expiry at a time, which is significantly easier than calibrating all expiries at once. Similarly to chapter 4, we are able to achieve an accurate fit to FTSE 100 index option prices, and again confirm the significant effect that interest rate stochasticity can have on the prices of long-dated barrier options. Furthermore, we demonstrated the accuracy of our simulation procedure by comparing the Monte-Carlo and Market prices of vanilla options.

Instead of fitting a specific parametric model to market data, and then using this model to price more exotic derivatives, it is sometimes possible to draw a direct link between the prices of liquid instruments and the derivative in question. This non-parametric approach has the key advantage that the resulting price will be valid for all models satisfying the underlying assumptions, and won't depend on the specific choice of parametrization or calibration techniques used. In the case of general volatility derivatives, existing model-free work has assumed that interest rates are deterministic. Although this is a safe assumption for short-term contracts, the prices of those with longer-terms may be significantly influenced by stochastic rates. Therefore, in chapter 6 we extended the work of Carr and Lee (2008) to include stochastic interest rates.

Conditional on a model for interest rates, we are able to write the prices of a variety of volatility derivatives in terms of the prices of power options, which are non-path-dependent and can therefore be replicated using vanilla calls and puts. Our main assumptions are that the underlying bond and stock price processes are continuous, and that their volatility processes and driving Brownian motions are all mutually independent. This independence assumption is partially relaxed by making our results first-order immune to the correlation between the stock's volatility process and its driving Brownian motion.

In order to provide an empirical example, it is necessary to fit both a model for interest rates and a curve to the market prices of vanilla option. In particular, we used a Hull-White short-rate model coupled with a mixture of log-normal distributions fit to Euro Stoxx 50 index options expiring in 9.15 years. Through this example we saw that the prices of a variety of volatility derivatives, including swaps, calls and puts on both volatility and variance, may have significantly different prices if
interest rates are assumed to be stochastic instead of deterministic.
References


References

Minerva Access is the Institutional Repository of The University of Melbourne

**Author/s:**
Ranasinghe, Navin

**Title:**
Pricing long-dated equity derivatives under stochastic interest rates

**Date:**
2017

**Persistent Link:**
http://hdl.handle.net/11343/145394

**File Description:**
Pricing long-dated equity derivatives under stochastic interest rates

**Terms and Conditions:**
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.