ABSTRACT

Title of dissertation:	PRICING VARIANCE DERIVATIVES USING HYBRID MODELS WITH STOCHASTIC INTEREST RATES
	Taras Smetaniouk, Doctor of Philosophy, 2008
Dissertation directed by:	Professor Dilip Madan Applied Mathematics and Scientific Computation University of Maryland, College Park

In this thesis, the research focuses on the development and implementation of two hybrid models for pricing variance swaps and variance options. Some variance derivatives (i.e., variance swap) are priced using portfolios of put and call options. However, longer-term options price not only stock variance, but also interest rate variance. By ignoring stochastic interest rates, variance derivatives utilizing this approach are overpriced. In recent months, the Federal Reserve lowered the funds rate as the equity markets fell. This created correlation between equities and interest rates. Furthermore, interest rate volatility increased. Thus, it is presently crucial to understand how stochastic interest rates and correlation impact the pricing of variance derivatives.

The first model (SR-LV) is driven by two processes: the stock return follows a diffusion and the stochastic interest rate is driven by the Hull-White [16] short rate dynamics. Local volatility is constructed with the help of Gyöngy's [14] result on recovering a Markov process from a general n-dimensional Itô process with the same

marginals. In the solution for the local volatility, the joint forward density of the stock price and interest rate is derived by solving an appropriate partial differential equation. Realized variance can then be computed by Monte Carlo simulation under the forward measure where local variances are collected over each realized path and averaged. Results are presented for different levels of assumed correlation between the stock price and interest rates. Prices obtained are lower than those produced with an options portfolio and this price difference strongly depends on the volatility of the short rate.

The second model (SR-SLV) adds one more dimension to the first model. In practice, volatility of a stock may change without the stock price moving. This effect is not captured in SR-LV model, but stochastic local volatility exhibits this trait. In this setting, a leverage function must be calibrated utilizing the joint density of the stock price, interest rate, and a stochastic term governed by a mean reverting lognormal model. By design, the price of variance swaps is the same as under SR-LV dynamics. However, variance option prices differ from SR-LV model and are presented for different levels parameters of the new stochastic component.

Although this work focuses on pricing variance derivatives, the developed methodology is extended to pricing volatility swaps and options.

PRICING VARIANCE DERIVATIVES USINGS HYBRID MODELS WITH STOCHASTIC INTEREST RATES

Taras Smetaniouk

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2008

Advisory Committee:
Professor Dilip Madan, Chair/Advisor
Professor Konstantina Trivisa
Professor C. David Levermore
Assistant Professor Anna Obizhaeva
Professor Victor M. Yakovenko

DEDICATION

To my wife, Olena, for her continuous love and support.

To our parents for the opportunities in life they created for us.

ACKNOWLEDGMENTS

Although at the moment of writing this thesis it may seem a lonely venture, there are many people to whom I owe a great deal of gratitude for being here at this stage of my life. First, I would like to thank my advisor, Dr. Dilip Madan. Although often nervous and unsure of myself when I entered his office, I always left with the greatest insight into the problem at hand when I exited. This sentiment has resonated among several students whom I know. He taught me to understand problems from a realistic point of view, not just mathematical. He always presented problems in his office and the math finance research interaction team (RIT) from a very practical point of view. Dr. Madan is a unique individual who has one foot in academia and the other in the industry. I greatly benefited from this by working on a relevant problem with interesting mathematical challenges.

Thank you to Dr. Michael Fu for organizing the weekly meeting for the mathematical finance RIT. Here, I was exposed to many different areas of mathematical finance and benefitted from presenting my own work. I was also a student in Dr. Fu's course in stochastic processes. This rigorous course gave me the tools needed to understand and implement stochastic processes in my own work.

I would like to thank Dr. Levermore who first provided me with the financial support for graduate school. I still remember his advise about writing an abstract, as it applied to a preliminary oral exam then. He said that in the first few sentences one should tell the reader why should he/she care about this work. The attention must be grabbed from the beginning.

Dr. James Yorke has also given me the opportunity to work at the Applied Mathematics and Scientific Computation department office. During this time, I learned how the department runs and acted as an advisor to students on various issues. So, thank you Dr. Yorke for this opportunity.

I would also like to that Dr. Jeffery Cooper who was my advisor when I entered the university. He also encouraged me to apply for the dissertation completion fellowship which I subsequently received.

I am grateful to Dr. Trivisa for the PDE course she taught me in my first year of graduate school. I had been away from school for a couple of years, so getting back into mathematics was a bit challenging at times. Her course was definitely challenging, but with her help I got through it and currently working actively with PDEs.

I have spoken to several individuals outside of the university regarding my research. I would like to thank Christopher Jordinson from Deutsche Bank Quantitative Products team for his thoughts on density PDE and simulation. Thanks to Daniel Duffy from Datasim for several discussion on splitting methods for PDEs with mixed derivatives. Also, special thanks to Lester Coyle of AVM LP., for the internship opportunity on the credit desk during the summer of 2007 where I was able to work with traders in one of the more dynamic and memorable months in credit derivatives history.

Thank you to my classmates for their help on different issues, especially Samvit

Prakash. Additionally, Mrs. Alverda McCoy of AMSC has been very helpful with any questions I had. I really enjoyed our talks on different matters and we always had a good laugh about things. So, thank you.

Thank you to all the committee members for taking their time to come to the defense and listen to my work.

It is likely that I forgot someone, so please forgive me. No man is an island. So, thank you all and I could not have done this alone.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Variance and volatility financial derivatives have recently gained popularity because they provide investors with pure exposure to the variance or volatility of some underlying asset. Historically traded over-the-counter (i.e., privately between companies), some volatility products are now traded in the market, such as the VIX options. Thus, there is a need to understand the fundamentals of how to price such derivatives correctly and efficiently.

This research is extremely relevant at present time. Currently, the economy is on the brink of a recession fueled by the credit crisis. Over the last several months, the Federal Reserve lowered the funds rate with the falling equity markets. This action created correlation between equities and the interest rates. Furthermore, the volatility of the interest rates increased. Therefore, it is crucial to account for correlation and stochasticity of the interest rates in the pricing of variance derivatives.

A variance swap is traditionally priced using an appropriately weighted portfolio of put and call options which mimics the terminal payoff of the swap. However, longer-term options price not only stock variance, but also interest rate variance. Thus, by ignoring stochastic interest rates, variance swaps are overpriced. The primary contribution of this research to mathematical finance is the development of two models for pricing variance derivatives correctly accounting for the interest rate variance. The first model includes stochastic interest rates and local volatility (SR-LV), while the second has a stochastic local volatility component (SR-SLV). In addition to more realistic volatility dynamics, the SR-SLV model prices options on realized variance.

Results presented are prices of variance swaps for different levels of assumed correlation between the stock price and interest rate for different maturities. Options on variance are priced at several correlation levels to examine the role of correlation with other factors, such as time and short rate volatility. The methodology is easily extended to price volatility derivatives, so a section in this dissertation is devoted to volatility swaps and options.

1.1 Variance swaps/options

In a variance swap contract, realized variance is defined over business days $0 < t_1 < ... < t_n = T$ for some stock, S, as

$$RV(0,T) = \frac{252}{n}\sum_{i=1}^{n}\left(\ln\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2. \tag{1.1}$$

The variance swap contract then has a payoff at maturity given by

$$Notional * (RV(0,T) - K),$$

where K is the agreed upon strike variance at initiation of the contract and the notional is the dollar amount on which the contract is written (e.g., $100 million). The value of the variance swap at time 0:

$$Notional * E\Big[e^{-\int_0^t r_s ds}(RV(0,T) - K)\Big].$$

The expectation is taken with respect to the risk-neutral probability measure where the money market account is the numeraire. Because this expectation is awkward to evaluate, it is easier to change measure to the T-forward measure, where the numeraire is a discount bond. Thus, the value of the variance swap is now given by

$$Notional * P(0,T) * E^{Q_T}\Big[(RV(0,T) - K)\Big],$$

where $P(0,T)$ is the price of the zero coupon bond maturing at time T and the expectation is taken with respect to the T-forward measure (i.e., notation E^{Q_T}).

The value of the variance swap must be zero at its initiation. Setting the above expression to zero gives the price (agreed upon fair variance), K_{vs}, of swap to

$$K_{vs} = E^{Q_T}\Big[RV(0,T)\Big]. \tag{1.2}$$

Therefore, the problem of pricing variance swaps reduces to evaluating this expectation.

Time 0 value of options on realized variance is given by:

$Notional * P(0,T) * E^{Q_T}\Big[(RV(0,T) - K)^+\Big]$ for a call option and

$Notional * P(0,T) * E^{Q_T}\Big[(K - RV(0,T))^+\Big]$ for a put option.

1.2 Models

The SR-LV model is given by the following system of SDEs:

$$\frac{dS_t}{S_t} = (r_t - q)dt + \sigma^s(S,t)dW_t^s,$$

$$dr_t = (\theta_t - kr_t)dt + \sigma^r dW_t^r,$$

$$dW_t^s dW_t^r = \rho dt.$$

The stock price return process is driven by a single Brownian motion, r_t is the stochastic interest rate, q is the dividend yield, and $\sigma^s(S,t)$ is the local volatility dependent on stock price and time. The interest rate process (Hull-White [16]) is driven by another Brownian motion with a long-term mean θ_t, a speed of mean reversion k, and a rate volatility σ^r. Correlation between the Brownian motions is set to a predetermined ρ. Realized variance in this setting can be calculated as

$$RV(0,T) = \frac{1}{T}\int_0^T \sigma^{s2}(S,t)dt.$$

Using Monte Carlo simulations, for a given path, the above integral can be approximated by a sum.

The SR-SLV model is given by the following system of SDEs:

$$\begin{aligned}
\frac{dS_t}{S_t} &= (r_t - q)dt + \sigma_*^s Z_t dW_t^s, \\
dr_t &= (\theta_t - kr_t)dt + \sigma^r dW_t^r, \\
d\ln Z_t &= \nu(\zeta_t - \ln Z_t)dt + \lambda dW_t^z, \\
dW_t^s dW_t^r &= \rho dt,
\end{aligned}$$

where dW_t^z is uncorrelated with dW_t^s and dW_t^r. The relationship between local volatility and the leverage function, σ_*^s, is $\sigma^{s2}(S,t) = \sigma_*^{s2} E[Z_t^2|S_t = S]$ (given by Gyöngy [14]). Realized variance in this case is given by:

$$\frac{1}{T}\int_0^T \sigma_*^{s2} Z_t^2 dt.$$

1.3 Local volatility

One of the main challenges in this work is recovering the local volatility function. Local volatility under deterministic interest rates is a well understood concept which was originally developed by Bruno Dupire [11]. In this setting, local volatility is given by the following deterministic function of strike and time:

$$\sigma^2(K,T) = \frac{\frac{\partial C}{\partial T} + qC + (r_T - q)K\frac{\partial C}{\partial K}}{\frac{K^2}{2}\frac{\partial^2 C}{\partial K^2}},$$

where C is the price of a call option with strike K and maturity T. Thus, given an options surface, the local volatility surface can easily be recovered.

Gyöngy [14] has shown how to construct a Markov process with the same marginals from some multi-dimensional Itô process. Local volatility under stochastic interest rates has recently been addressed by Atlan [1] and Ren *et al.* [19] who applied Gyöngy's [14] result to arrive at the new form for local volatility:

$$\sigma^2(K,T) = \frac{\frac{\partial C}{\partial T} + q(C - KC_K) + KE\Big[\exp(-\int_0^T r_t dt) r_T \mathbf{1}_{S_T > K}\Big]}{\frac{K^2}{2}\frac{\partial^2 C}{\partial K^2}}. \tag{1.3}$$

The expectation in this expression is given under the risk-neutral measure. Again, changing to T-forward measure simplifies the work, where

$$E\Big[\exp(-\int_0^T r_t dt) r_T \mathbf{1}_{S_T > K}\Big] = P(0,T) E^{Q_T}\Big[r_T \mathbf{1}_{S_T > K}\Big]. \tag{1.4}$$

To evaluate the expectation on the right-hand side, the joint forward density must be know. That is, for SR-LV model, this is a two dimensional density in stock and interest rate. Therefore, solving for local volatilities for all times is a consequence of knowing the joint densities at all times. This problem is tackled by solving an appropriate partial differential equation (PDE) forward one step at a time and recovering local volatilities.

For the SR-SLV model, local volatility is an input from SR-LV model. Keeping terminology consistent with Ren *et al.* [19], the leverage function, σ_*^s, must be calibrated using local volatilities and $E[Z_t^2|S_t = K]$. Thus, the joint distribution of stock, interest rate, and Z must be known at each time. In this setting, stepping through the solution of a PDE in three variable is required to iteratively recover the leverage function.

In both models, the PDEs contain cross derivative terms and traditional finite difference schemes do not perform well. An ADI scheme by Craig and Sneyd [6] is implemented for the two PDE solvers to address this problem.

1.4 Data

Results are provided for data obtained for S&P 500 index options, yield curve, interest rate swaptions, and caps for June 19, 2002. The Hull-White short rate model must be calibrated to the yield curve, swaptions, and caps simultaneously. The short rate process is a Gaussian process, so there exists a closed form for the price of a discount bond which may be utilized in the calibration. Since swaptions and caps are easily priced on trees, Hull's [17] calibration methods is chosen and the details are provided in the Appendix. From this time forward, assume that the calibrated parameters for the Hull-White short rate model are $k = 0.025091$ and

$\sigma^r = 0.011591$.

On the equity side, an options surface is needed for any strike and maturity up to three years to study the impact of stochastic rates on variance derivatives. Market data is available for only certain strikes and maturities, thus a model needs to be fit to the data. In general, one cannot fit a single set of parameters in a model and calibrate across time. Recently, Carr *et al.* [4], developed the theory for self-decomposable laws and associated processes. This result can be applied to the Variance Gamma process for the evolution of the stock price developed by Madan and Seneta [20]. The resulting VGSSD framework allows calibration to the market options surface of a single set of parameters. Then, an arbitrarily fine grid of options prices can be generated. The Variance Gamma process was chosen because the corresponding distribution fits market data well; that is, its parameters control additional aspects of skewness and kurtosis. Furthermore, Fast Fourier Transform can be employed in generating option prices which significantly speeds up calibration.

One problem to consider is that the options data available for S&P 500 is of maturities up to 2 years and if considering volume of options traded at the later maturities, then only maturities of up to 1.5 years are available for calibration. One solution applied here is the creation of a stylized options surface where VGSSD parameters are calibrated from existing options, but the time scaling parameter (γ) is increased to compensate for lack of longer dated options. In practice, financial institutions have this options surface for all strikes and maturities, so this adjustment is a remedy for the lack of data. Details of implementation of VGSSD are in the Appendix. From this moment, assume that the options surface can be generated via

VGSSD with for parameters: $\theta = -0.16782, \nu = 0.65289, \sigma = 0.19732$, and $\gamma = 0.6$.

1.5 Monte Carlo simulation

For both models, variance swap and option prices are determined by Monte Carlo simulation. Gyöngy's [14] result is utilized in SR-LV model by simulating a single stock price process in the forward measure using the densities derived. The same approach can be taken with the SR-SLV when pricing variance swaps. But, by construction, both models give the same price for variance swaps.

All three processes need to be simulated for options on realized variance in the forward measure. In SR-SLV model, local volatility is recovered and then the leverage function is calibrated. This involves two PDE solvers, so computational time needs to be considered. Taking larger time steps requires more simulations for accuracy. So, for this model a second order accurate scheme is considered for simulation, illustrated in Glasserman [13]. However, this procedure involves first and second partials of the leverage function with respect to the stock. For stability, these partials need to be bounded artificially. Ultimately, a simpler discretization is chosen since increasing the number of simulations minimally impacted computation time.

The two models confirmed the hypothesis that pricing by a replicating portfolio of options overprices variance swaps (with a few later-discussed exceptions). Option prices, generated by the SR-SLV model, increase as a function of vol of vol and decrease as a function of speed of mean reversion of Z as expected. Further analysis of swaps and options is conducted across different correlations and maturities. These results are summarized in Chapter 4.

Chapter 2

Traditional Pricing of Variance Swaps

Log-contract hedge is one common way to price variance swaps. This approach involves a replicating portfolio of out-of-the-money call and put options on the underlying stock. In theory, the continuum of strikes is needed for a given time, but an approximation can be obtained with finite number of strikes. This approach is outlined in Derman *el al.* [7] and Bossu and Strasser [**?**].

Since the option surface in this project is generated using a VGSSD model, it is much easier to price a variance swap because for this model, the characteristic function for $\ln S_T$ is known in closed form exactly. This approach is described by Schoutens [27]. In this chapter, both methods are reviewed and produce the same price under the assumption of constant interest rates. Prices generated for variance swaps in this chapter will be used as a baseline for comparison when the interest rate becomes stochastic. Schoutens [27] scheme is preferred since options surface is generated by a VGSSD model.

2.1 Pricing by log-contract hedge

In the first chapter, it was shown that the realized variance of stock S_t from time 0 to T is given by

$$RV(0,T) = \frac{252}{n}\sum_{i=1}^{n}\left(\ln\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2. \tag{2.1}$$

From the Taylor approximation we get the following:

$$\sum_{i=1}^{n}\left(\ln\frac{S_{t_i}}{S_{t_{i-1}}}\right)^2 \approx \sum_{i=1}^{n}\left(\frac{S_{t_i}-S_{t_{i-1}}}{S_{t_{i-1}}}\right)^2 \approx 2\left[\sum_{i=1}^{n}\left(\frac{S_{t_i}-S_{t_{i-1}}}{S_{t_{i-1}}}\right) - \ln\frac{S_T}{S_0}\right]. \tag{2.2}$$

Continuously-sampled realized variance can then be written as

$$RV(0,T) = \frac{2}{T}\left[\int_0^T \frac{dS_t}{S_t} - \ln\frac{S_T}{S_0}\right]. \tag{2.3}$$

From the above expression, buying realized variance is equivalent to a long dynamic position in the stock and a short static position in the log contract. Taking the risk-neutral expectation yield the price of the variance swap:

$$K_{vs} = \frac{2}{T}E\left[\int_0^T \frac{dS_t}{S_t} - \ln\frac{S_T}{S_0}\right] = \frac{2}{T}\left[(r-q)T - E\ln\frac{S_T}{S_0}\right], \tag{2.4}$$

where r is the risk-free interest rate for time T, q is the continuous dividend yield, and $(r-q)$ is the expected growth rate of this dividend paying stock.

Derman *et al.* [7] shows how to replicate the payoff of the log contract over some reference stock price S_* (e.g., S_0 or forward strike), with out-of-the-money calls and puts. First, the log payoff can be written as

$$\ln\frac{S_T}{S_0} = \ln\frac{S_T}{S_*} + \ln\frac{S_*}{S_0}, \tag{2.5}$$

where $\ln\frac{S_T}{S_*}$ can be further decomposed into a long position in $(1/S_*)$ forward contracts with strike S_*, short position in $(1/K^2)$ put options with strike K for all strikes from 0 to S_*, and short position in $(1/K^2)$ call options with strike K for all strikes:

$$\ln\frac{S_T}{S_*} = \frac{S_T - S_*}{S_*} - \int_0^{S_*}\frac{1}{K^2}(K-S_T)^+dK - \int_0^{S_*}\frac{1}{K^2}(S_T-K)^+dK. \tag{2.6}$$

Substitution Equations 2.5 and 2.6 into 2.4 and taking expectations gives

$$K_{vs} = \frac{2}{T}\left[(r-q)T - \left(\frac{S_0}{S_*}e^{(r-q)T}\right) - \ln\frac{S_*}{S_0}\right] + e^{rT}\Pi_{CP}, \tag{2.7}$$

where Π_{CP} is the present value of the portfolio of options with payoff at expiry given by

$$f(S_T) = \frac{2}{T}\left[\frac{S_T - S_*}{S_*} - \ln\frac{S_T}{S_*}\right]. \tag{2.8}$$

In reality, only a finite number of options are available for certain strikes. Given this limitation, $f(S_T)$ can be approximated, for example, by a piece-wise linear function. In practice, this linear approximation will overestimate the value of this payoff. However, in this dissertation, options surface is generated via model (VGSSD) and option values can be generated for any strike desired. Thus, an optimal number of strikes can be chosen for desired accuracy of pricing. Computational time is not a concern, since in a VG setting fast Fourier transform (FFT) [5] is utilized in options pricing that greatly reduces computational time.

Derman *et al* [7] proposed a lined approximation and his notation is followed. Assume that the following strikes are trading for calls and puts in the market:

$$...K_{3p} < K_{2p} < K_{1p} < K_0 = S_* < K_{1c} < K_{2c} < K_{3c}...$$

Then, Π_{CP} can be calculated by the following:

$$\Pi_{CP} = \sum_i w_p(K_{ip})P(S, K_{ip}) + \sum_i w_c(K_{ic})C(S, K_{ic}),$$

where

$$w_c(K_{n,c}) = \frac{f(K_{n+1,c}) - f(K_{n,c})}{K_{n+1,c} - K_{n,c}} - \sum_{i=0}^{n-1} w_c(K_{i,c}),$$

$$w_p(K_{n,p}) = \frac{f(K_{n,p}) - f(K_{n+1,p})}{K_{n,p} - K_{n+1,p}} - \sum_{i=0}^{n-1} w_p(K_{i,p}).$$

Functions $P(S,K)$ and $C(S,K)$ represent the prices of put and call options respectively. The weights, $w(K)$, are the number of options needed to be held at strike K.

2.1.1 Example

In this example, a one year variance swap on the S&P 500 index for June 19, 2002 will be priced using the above methodology. On this day, the S&P 500 index closed at 1020, the one year risk-free rate is 0.02359, and the dividend yield is .0118. Consider the range of strikes for this exercise to be from 20% to 200% of the close of the index at increments of 100. The reference stock price, $S_* = 1002$, the one closest to the one-year forward. Calibrated VGSSD parameters for the options surface are $\theta = -0.16782, \nu = 0.65289, \sigma = 0.19732$, and $\gamma = 0.6$. The replicating portfolio is given in Table 2.1.

From notation in this section, Π_{CP} is 539.476. With these options the fair strike (price) for the variance swap is give by Equation 2.7 and is equal to $(23.502)^2$. It is common to quote variance swap price in terms of volatility, and it is understood that the actual price is volatility squared. Thus, from this moment, prices will be quoted in volatility (e.g., from above - 23.502).

Since a piece-wise linear approximation is used for convex payoff $f(S_T)$, the calculated fair variance is too high. But if more options are used at closer strikes, then in the limit we obtain the correct variance swap price. If the strikes in this

Put/Call	Strike	Option Price	Weight	Contribution
P	202	0.004	56.227	0.230
P	302	0.051	23.227	1.189
P	402	0.303	12.775	3.876
P	502	1.193	8.098	9.658
P	602	3.607	5.596	20.186
P	702	9.088	4.100	37.264
P	802	19.992	3.134	62.652
P	902	39.541	2.473	97.802
P	1002	71.673	1.068	76.523
C	1002	101.070	0.934	94.432
C	1102	52.192	1.654	86.310
C	1202	22.685	1.389	31.512
C	1302	9.391	1.183	11.112
C	1402	3.988	1.020	4.068
C	1502	1.761	0.888	1.564
C	1602	0.809	0.781	0.632
C	1702	0.387	0.692	0.268
C	1802	0.192	0.617	0.118
C	1902	0.098	0.554	0.054
C	2002	0.052	0.500	0.026
			TOTAL	539.476

Table 2.1: Options portfolio for S&P 500 options on June 19, 2002 which replicate payoff given by $f(S_T)$.

example are \$50 apart, the price becomes 23.082 and if \$10 apart - 22.903. Since options prices are generated by a model where the characteristic function for $\ln(S_T)$ is known, the next section will show the limit price to be 22.9.

2.2 Pricing by characteristic function

Schoutens [27] proposes a simpler approach for pricing the log contract. When the characteristic function for $\ln S_T$ is known in closed from, it is straightforward to evaluate $E[\ln S_T]$. Suppose the risk-neutral characteristic function of $\ln S_T$ is

$$\phi(u,t) = E\Big[\exp(iu\ln(S_T))\Big].$$

Then

$$-i\frac{\partial\phi(u,t)}{\partial u}\bigg|_{u=0} = E\Big[\ln(S_T)\Big].$$

The price of variance swap, K_{vs}, from Equation 2.4 becomes

$$K_{vs} = \frac{2}{T}\left[(r-q)T - E\ln\frac{S_T}{S_0}\right] = \frac{2}{T}\left[(r-q)T + i\frac{\partial\phi(0,t)}{\partial u} + \ln S_0\right]. \tag{2.9}$$

For VGSSD the characteristic function for $\ln S_T$ is given in Carr *et al.* [4] as

$$E\Big[\exp(iu\ln(S_T))\Big] = \exp\Big[iu(\ln(S_0) + (r-q)t - \ln\psi(-i))\Big]\ln\psi(u),$$

where

$$\psi(u,t) = \left(\frac{1}{1 - iu\theta\nu t^{\gamma} + \frac{\sigma^2\nu}{2}u^2t^{2\gamma}}\right)^{1/\nu}$$

More details on VGSSD can be found in the Appendix.

Using model inputs from the previous example, variance swaps can now be priced for any given time. Table 2.2 shows the term structure for variance swap prices (in volatility). Note that the one year variance swap price is 22.9, which is the limit of Example 2.1.1 when the number of options increased and the distance between strikes decreased.

The shape in time for the prices in Table 2.2 is consistent with observed variance swap curves in literature (i.e.,increasing in time, but at a decreasing rate). These prices will serve as a basis of comparison for the later developed models. The hypothesis in this work is that the prices in Table 2.2 are too high because they do not account for the stochasticity of interest rates.

Expiry (yrs)	1	1.5	2	2.5	3	3.5	4
K_{vs} in vol	22.900	23.589	24.059	24.410	24.686	24.913	25.103

Table 2.2: Term structure of variance swaps for Example 2.1.1 using VGSSD characteristic function for $\ln S_T$

It is worth noting that the characteristic function for $\ln S_T$ in practice is often unknown. Pricing via characteristic function in this research is the benefit of working with a stylistic options surface generated by VGSSD. For the models developed later, there is no closed form solution for the characteristic function and thus simulation is used for pricing.

Chapter 3

Local volatility and Stochastic Local Volatility via PDEs

This chapter provides details for generating the local volatility surface and the leverage function for the two models: Stochastic Rates - Local Volatility (SR-LV) and Stochastic Rate - Stochastic Local Volatility (SR-SLV), respectively. The derived expression for local volatility is a function of an expectation under the T-forward measure. For the SR-LV model, the solution to the PDE for the joint density of interest rate and stock is utilized in evaluating this expectation for any strike and time. Then, a leverage function must be calibrated to the local volatility surface for the SR-SLV model, where another PDE is solved in three dimensions to recover this function.

3.1 Gyöngy's result

In the section, Gyöngy's [14] result is summarized and applications discussed. Let $\xi(t)$ be a stochastic process (one or multi-dimensional) starting from 0 with Itô differential

$$d\xi(t) = \beta(t, \omega)dt + \delta(t, \omega)dW(t),$$

where $W(t)$ is a Wiener process, δ and β are bounded. For every time t, this process has marginal distributions for the random variable(s) $\xi(t)$. Suppose we want to construct a Markov process $x(t)$ that has the same marginals as $\xi(t)$. Gyöngy not only proved the existence but provided the machinery on how to construct such a process. So, let $x(t)$ be given by

$$dx(t) = b(t, x(t))dt + \sigma(t, x(t))dW(t),$$

then Gyöngy tells us that

$$b(t, x(t)) = E\Big[\beta(t,\omega)|\xi(t) = x\Big],$$
$$\sigma^2(t, x(t)) = E\Big[\delta(t,\omega)\delta^T(t,\omega)|\xi(t) = x\Big].$$

This result will be utilized in construction of the leverage function for the SR-SLV model and in simulating the models.

3.2 Local volatility under stochastic interest rates

The following derivation is found in Madan *et al.* [19], but the same expression for local volatility is also derived in Atlan [1]. Assume the following dynamics for the stock involving stochastic interest rates:

$$\frac{dS_t}{S_t} = (r_t - q)dt + \sigma(S_t, t)dW_t^S.$$

The call option price can be written as

$$C(K, T) = E\left[e^{-\int_0^T r_u du}(S_T - K)^+\right]$$

and

$$C_K(K, T) = E\left[e^{-\int_0^T r_u du}\mathbf{1}_{S_T > K}\right],$$
$$C_{KK}(K, T) = E\left[e^{-\int_0^T r_u du}\delta_{S_u = K}\right],$$
$$C(K, T) - KC_K(K, T) = E\left[e^{-\int_0^T r_u du}S_T\mathbf{1}_{S_T > K}\right].$$

Using Itô, the final discounted payoff has the following form:

$$e^{-\int_0^T r_u du}(S_T - K)^+ = (S_0 - K)^+ - \int_0^T du\ r_u e^{-\int_0^u r_v dv}(S_u - K)^+ -$$

$$\int_0^T du\ e^{-\int_0^u r_v dv}\mathbf{1}_{S_u>K} dS_u + \frac{1}{2}\int_0^T du\ e^{-\int_0^u r_v dv}\delta_{S_u=K}\ \sigma^2(S_u, u)S_u^2 du.$$

The call price is then obtained by taking expectations:

$$C(K,T) = (S_0 - K)^+ - E\left[\int_0^T du\ r_u e^{-\int_0^u r_v dv}(S_u - K)^+\right] -$$

$$\int_0^T du\ E\left[e^{-\int_0^u r_v dv}\mathbf{1}_{S_u>K}(r_u - q)S_u\right] + \frac{1}{2}\int_0^T du\ \sigma^2(K,u)K^2 E\left[e^{-\int_0^u r_v dv}\delta_{S_u=K}\right].$$

Differentiating both sides with respect to T, we get

$$C_T(K,T) = -E\left[\ r_T e^{-\int_0^T r_v dv}(S_T - K)^+\right] -$$

$$E\left[e^{-\int_0^T r_v dv}\mathbf{1}_{S_T>K}(r_T - q)S_T\right] + \frac{1}{2}\sigma^2(K,T)K^2 E\left[e^{-\int_0^T r_v dv}\delta_{S_T=K}\right].$$

Using the above established identities, this simplifies to

$$C_T(K,T) = -E\left[\ r_T e^{-\int_0^T r_v dv}(S_T - K)\mathbf{1}_{S_T>K}\right] - E\left[r_T e^{-\int_0^T r_v dv}\mathbf{1}_{S_T>K}S_T\right] -$$

$$q(C(K,T) - KC_K(K,T)) + \frac{1}{2}\sigma^2(K,T)K^2 C_{KK}(K,T).$$

Simplifying yields

$$C_T(K,T) = KE\left[\ r_T e^{-\int_0^T r_v dv}\mathbf{1}_{S_T>K}\right] - q(C(K,T) - KC_K(K,T)) +$$

$$\frac{1}{2}\sigma^2(K,T)K^2 C_{KK}(K,T).$$

Finally, solving for local variance,

$$\sigma^2(K,T) = 2\frac{C_T + q(C - KC_K) - KE\left[e^{-\int_0^T r_v dv} r_T \mathbf{1}_{S_T>K}\right]}{K^2 C_{KK}}. \tag{3.1}$$

In practice, the expectation in equation 3.1 is difficult to evaluate. So, one approach is to work under a different measure. Under the risk-neutral measure, the numeraire is the money market account. While, under the T-forward measure, the numeraire is the discount bond, $P(0,T)$. The Radon-Nikodym derivative for the measure change is just the ratio of the two numeraires. Thus, the expectation in equation 3.1 can be written under the T-forward measure as

$$E\left[e^{-\int_0^T r_v dv} r_T \mathbf{1}_{S_T>K}\right] = P(0,T)E^{Q_T}\left[r_T \mathbf{1}_{S_T>K}\right] \tag{3.2}$$

Now, one can consider $r_T \mathbf{1}_{S_T>K}$ as a terminal payoff for some derivative and evaluate the expectation by solving an appropriate PDE. But, suppose that the local volatility surface lives on a range of 50 strikes. This means that, for each time t, there is a need to solve 50 PDEs. Although solving one is quick, this is not practical.

A better approach is to solve one PDE for the joint density of interest rates and stock. Then, this density can be used to numerically evaluate the expectation in equation 3.2 for all K. This is the motivation for building and solving a PDE for the joint density in this chapter.

3.3 Forward PDE for the joint density in a multi-factor setting

Overhaus et al. [23], Chapter 8, provides an elegant derivation of the forward PDE for the t-forward density from risk-neutral processes. Suppose that for $1 \leq i \leq n$, there are some general risk-neutral processes:

$$dx_i = \mu_i(\mathbf{x}_t, t)dt + \sigma_i(\mathbf{x}_t, t)dW_t^i,$$

$$d\langle W_t^i, W_t^j \rangle = \rho_{ij} dt.$$

Let $V(\mathbf{x}, t)$ represent the price of a derivative and the money market account is $B_t = e^{\int_0^t r_s ds}$. In the risk-neutral setting, V/B is a martingale. Using Itô's lemma, one can write a PDE for V/B and by setting the drift equal to zero to arrive at

$$\frac{\partial V}{\partial t} - r_t V + \sum_i \mu_i \frac{\partial V}{\partial x_i} + \frac{1}{2} \sum_{ij} \rho_{ij} \sigma_i \sigma_j \frac{\partial^2 V}{\partial x_i \partial x_j} = 0. \tag{3.3}$$

Arrow-Debreu prices and the t-forward measure are related in the following identity:

$$\psi(\mathbf{x}, t) = P(0, t) \phi(\mathbf{x}, t), \tag{3.4}$$

where $\psi(\mathbf{x}, t)$ is the Arrow-Debreu price representing the value of a derivative with a \$1 payoff in state $\mathbf{x}_t$, $\phi(\mathbf{x}, t)$ is the t-forward density, and $P(0, t)$ is the discount bond price expiring at time t.

The relationship in Equation 3.4 comes from the change of numeraire between the risk-neutral and t-forward measures. The numeraire under the risk-neutral measure is the money market account, while under the t-forward measure it is $P(0, t)$. That is,

$$V(\mathbf{x}_0, t) = \int V(\mathbf{x}, t) \psi(\mathbf{x}, t) d\mathbf{x} = P(0, t) \int V(\mathbf{x}, t) \phi(\mathbf{x}, t) d\mathbf{x}. \tag{3.5}$$

Differentiating both sides with respect to t, we get

$$0 = \int V \frac{\partial \psi}{\partial t} + \psi \frac{\partial V}{\partial t} d\mathbf{x}.$$

Substituting Equation 3.3 results in

$$0 = \int V \frac{\partial \psi}{\partial t} + \psi \left(r_t V - \sum_i \mu_i \frac{\partial V}{\partial x_i} - \frac{1}{2} \sum_{ij} \rho_{ij} \sigma_i \sigma_j \frac{\partial^2 V}{\partial x_i \partial x_j} \right) d\mathbf{x}.$$

Using integration by parts,

$$0 = \int V\left(\frac{\partial\psi}{\partial t} + \psi r_t + \sum_i \frac{\partial\mu_i\psi}{\partial x_i} - \frac{1}{2}\sum_{ij}\frac{\partial^2(\rho_{ij}\sigma_i\sigma_j\psi)}{\partial x_i\partial x_j}\right)d\mathbf{x} + \textit{boundary terms}.$$

Those boundary terms are well behaved functions of μ_i and σ_i, so they can be ignored. In order for the above equation to hold for any payoff $V(\mathbf{x}, t)$, it must be true that

$$0 = \frac{\partial\psi}{\partial t} + \psi r_t + \sum_i \frac{\partial\mu_i\psi}{\partial x_i} - \frac{1}{2}\sum_{ij}\frac{\partial^2(\rho_{ij}\sigma_i\sigma_j\psi)}{\partial x_i\partial x_j}.$$

This is the PDE for the density of the Arrow-Debreu price. The PDE for the t-forward density can be obtained using Equation 3.4 remembering the relationship between forward rate and bond prices

$$f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T},$$

where $f(t,T)$ is the rate determined at t for instantaneous borrowing at time T and $P(t,T)$ is the price of a discount bond at time t, maturing at T. The PDE for the t-forward is

$$0 = \frac{\partial\phi}{\partial t} + (r_t - f(0,t))\phi + \sum_i \frac{\partial\mu_i\phi}{\partial x_i} - \frac{1}{2}\sum_{ij}\frac{\partial^2(\rho_{ij}\sigma_i\sigma_j\phi)}{\partial x_i\partial x_j}. \tag{3.6}$$

It should be noted that for Hull-White/Vasicek interest rate model, $(r_t - f(0,t))$ is continuous. That is why at a later time we will work with this function.

In practice, the PDE given in Equation 3.6 is difficult to solve for two reasons. First, the initial condition is a delta function. Second, the PDE includes mixed derivative terms which are not easily handled by conventional finite difference methods. The particular PDE solver is the topic of the next section.

To solve the first problem, Overhaus *et al.* [23] suggests a change of variables. In this dissertation, there will be three dimensions at most, so consider Equation 3.6 in only three variables (x,y,z). Then, let

$$x = x' a_t,$$

$$y = y' b_t,$$

$$z = z' c_t,$$

where a_t, b_t, and c_t scale as $\sqrt{t}$ as $t \to 0$. This scaling is chosen because the distribution spreads out as $\sqrt{t}$. Then, the new variables - x',y', and z' - have the interpretation as the number of standard deviations away from the mean if a_t, b_t, and c_t are chosen to be the standard deviations for the marginal distributions for the respective processes. The density is also rescaled as

$$\phi(x, y, z) = \frac{\phi'(x', y', z')}{a_t b_t c_t},$$

and the following will hold:

$$\int \phi(x, y, z) dx dy dz = \int \phi'(x', y', z') dx' dy' dz'.$$

This change of variables allows for the initial distribution for ϕ' to be approximately normal, but very high peaked of course. The grid for the PDE must be defined far enough (e.g., in practice 4 or 5 standard deviations) for the boundary conditions to be zero everywhere. There is difficulty with discontinuities at $t = 0$, so the PDE starts at very small time away from zero (e.g., 10^{-6}). The details for the new PDEs for each of the two models are described in later sections of this chapter.

3.4 PDE solver

Finite difference methods are chosen for solving the PDEs for the two models. However, it is well known that standard finite difference methods do not handle well the mixed derivative terms implicitly [10]. Splitting or fractional steps approaches have been developed for tackling the mixed derivative problem (e.g., Yanenko [31], Craig and Sneyd [6], McKee and Mitchell [22]). That is, the N-dimensional problem is reduced to solving N tridiagonal systems.

Here is a simple example of splitting from Duffy [10]. Consider the two-dimensional heat equation:

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}.$$

Then the simple splitting procedure can be defined using explicit Euler

$$\frac{\tilde{U}_{ij} - U_{ij}^n}{\Delta t} = \Delta_x^2 U_{ij}^n,$$
$$\frac{\tilde{U}_{ij}^{n+1} - \tilde{U}_{ij}}{\Delta t} = \Delta_y^2 \tilde{U}_{ij},$$

where Δ_x^2 and Δ_y^2 are approximations of the second derivative using the central differences. A similar scheme can be defined implicitly,

$$\frac{\tilde{U}_{ij} - U_{ij}^n}{\Delta t} = \Delta_x^2 \tilde{U}_{ij},$$
$$\frac{\tilde{U}_{ij}^{n+1} - \tilde{U}_{ij}}{\Delta t} = \Delta_y^2 U_{ij}^{n+1}.$$

Assuming that the mesh size in both space directions is equal to h Duffy [10] shows that the explicit scheme is stable if $\Delta t/h^2 \leq .5$. The implicit scheme is unconditionally stable. ADI schemes do not exhibit unconditional stability.

In the previous example, there is no mixed derivative term. Craig and Sneyd [6] developed an ADI method and established stability conditions for problems with mixed derivatives. Suppose we have an initial value problem:

$$\frac{\partial u}{\partial t} = Lu, \tag{3.7}$$

where L is an elliptic partial differential operator

$$L = \sum_{i=1}^{N}\sum_{j=1}^{N} q_{ij}\partial_i\partial_j \qquad \partial_i = \frac{\partial}{\partial x_i}, \tag{3.8}$$

and whose coefficient may be functions of the x_i and t. The PDE is parabolic if the symmetric matrix $Q = (q_{ij})$ is positive definite.

Let $u^n_{j_1}, ..., u^n_{j_N}$ denote the finite difference solution at nodes $j_1\Delta x_1, ..., j_N\Delta x_N, n\Delta t$. Let $\Delta x_i = \Delta$, defining a uniform mesh. Also, define the following operations using central differences (example in two dimensions):

$$\delta_x^2 u_{ij} = u_{i+1,j} - 2u_{ij} + u_{i-1,j},$$

and

$$\delta_{xy} u_{ij} = u_{i+1,j+1} - u_{i+1,j-1} - u_{i-1,j+1} + u_{i-1,j-1}.$$

On this uniform mesh, define $r = \Delta t/\Delta^2$.

The Craig and Sneyd [6] "unsplit" scheme can be written in the following form:

$$Au^{n+1} = (A + B)u^n,$$

where

$$A = \prod_{i=1}^{N}(1 - \theta r q_{ii}\delta_{x_i}^2)$$

and

$$B = r \sum_{i=1}^{N} q_{ii} \delta_{x_i}^2 + \frac{1}{2} r \sum_{i=2}^{N} \sum_{j=1}^{i-1} q_{ij} \delta_{x_i x_j}.$$

Parameter θ determines the implicitness of the method and stability. That is, $\theta = 0$ corresponds to the fully explicit scheme and $\theta = 1$ corresponds to a fully implicit scheme, and stability will be addressed shortly. In implementation, this scheme is split into N tridiagonal matrix operations:

$$(1 - \theta q_{11} r \delta_{x_1}^2) u^{n+1(1)} = \left[1 + r(1-\theta) q_{11} \delta_{x_i}^2 + r \sum_{i=1}^{N} q_{ii} \delta_{x_i}^2 + \frac{1}{2} r \sum_{i=2}^{N} \sum_{j=1}^{i-1} q_{ij} \delta_{x_i x_j} \right] u^n,$$

$$(1 - \theta q_{22} r \delta_{x_2}^2) u^{n+1(2)} = u^{n+1(1)} - \theta q_{22} r \delta_{x_2}^2 u^n,$$

.

.

$$(1 - \theta q_{NN} r \delta_{x_N}^2) u^{n+1(N)} = u^{n+1(N-1)} - \theta q_{NN} r \delta_{x_N}^2 u^n,$$

where $u^{n+1(i)}$ is the approximation of u^{n+1} at split level (i). This setup includes other schemes as special cases. For example, when $\theta = 1$ and mixed derivative terms are absent, this is an unconditionally stable Douglas-Rachford [9] method for $N = 3$. Craig and Sneyd method is also a special case of McKee and Mitchell [22] with mixed derivatives.

Stability of the Craig and Sneyd scheme is only discussed in the constant coefficient (q_{ij}) case. Necessary and sufficient condition for stability in two dimensions is determined to be $\theta > 1/2$ and for three dimensions is $\theta > 2/3$. However, McKee and Mitchell [22] remark that the above conditions are necessary conditions for unconditional stability, but not sufficient since values above the indicated levels of θ have been found such that the scheme breaks down.

The problem in this dissertation work is that coefficients, q_{ij}, are not constant. Furthermore, these coefficients are unknown ahead of time for all t at the beginning of solving the PDE. The coefficients are recovered one step at a time walking through the PDE. To ensure stability in this work, a conservative approach was taken where above conditions on θ were satisfied and $\Delta t/\Delta^2 \leq .5$ was met. Stability was observed by ensuring that the probability density solution integrated to 100% (with error no more that 1% for any give time t) and the probabilities must be positive.

The discussion so far has been centered around the problem defined in Equations 3.7 and 3.8. In this dissertation work, the problem at hand has the following form:

$$\frac{\partial u}{\partial t} = \left(L + \sum_{i}^{N} q_i \partial_i + q \right) u, \tag{3.9}$$

where L is again an elliptic partial differential operator

$$L = \sum_{i=1}^{N} \sum_{j=1}^{N} q_{ij} \partial_i \partial_j \qquad \partial_i = \frac{\partial}{\partial x_i}. \tag{3.10}$$

However, the methodology described in this section is extended easily to this case and stability conditions remain unchanged for small values of Δt.

3.5 SR-LV model

This section discusses the recovery of the local volatility surface utilizing the joint density solution from an appropriate PDE. In the stochastic interest rates and local volatility framework, the SR-LV model is given by the following set of stochastic differential equations:

$$\frac{dS_t}{S_t} = (r_t - q)dt + \sigma^s(S,t)dW_t^s, \tag{3.11}$$

$$dr_t = (\theta_t - kr_t)dt + \sigma^r dW_t^r, \tag{3.12}$$

$$dW_t^s dW_t^r = \rho dt. \tag{3.13}$$

The stock price return process is driven by a single Brownian motion, r_t is the stochastic interest rate, q is the dividend yield, and $\sigma^s(S,t)$ (or just σ^s) is the local volatility dependent on stock price and time. The interest rate process (Hull-White [16]) is driven by another Brownian motion with a long-term mean θ_t, a speed of mean reversion k, and a rate volatility σ^r. Correlation between the Brownian motions is set to a predetermined ρ.

3.5.1 Hull-White as an O-U process

The Hull-White short rate process in Equation 3.12 can be rewritten as an Ornstein-Uhlenbeck process as shown in Overhaus *et al.* [23]:

$$dy_t = -ky_t dt + \sigma^r dW_t^r, \tag{3.14}$$

where now

$$r_t = y_t + \bar{y}_t$$

and

$$d\bar{y}_t = (\theta_t - k\bar{y}_t)dt.$$

The section in the Appendix on calibrating the Hull-White model only gives the values for the two parameters k and σ^r. Overhaus *et al.* [23] briefly shows why

θ_t does not come into the pricing of bonds and other derivatives in this setting and we can just work with the O-U process.

First, keeping the notation from Overhaus *et al.* define

$$\Lambda_t^s = \int_t^s k du = k(s-t), \qquad \hat{B}(t,T) = \int_t^T \exp(-\Lambda_t^s) ds = \frac{1-e^{-k(T-t)}}{k}.$$

Here is how $\bar{x}_t$ relates to the yield curve through discount bond price $P(t,T)$:

$$P(t,T) = E\Big[\exp\Big(-\int_t^T r_s ds\Big)|F_t\Big] = \exp\Big(-\int_t^T \bar{y}_s ds\Big)E\Big[\exp\Big(-\int_t^T y_s ds\Big)|F_t\Big]$$

Using the solution to Equation 3.14 from the Appendix,

$$P(t,T) = \exp\Big(-\int_t^T \bar{y}_s ds - y_t\hat{B}(t,T)\Big)E\Big[\exp\Big(-\sigma^r\int_t^T \hat{B}(s,T)dW_s\Big)|F_t\Big].$$

Simplifying results in

$$P(t,T) = \exp\Big(-\int_t^T \bar{y}_s ds - y_t\hat{B}(t,T) + \frac{1}{2}\sigma^{r2}\int_t^T \hat{B}(s,T)^2 ds\Big). \tag{3.15}$$

Letting $t = 0$, talking log of both sides of Equation 3.15 and differentiating with respect to T yields

$$f(0,T) = \bar{y}_T - \sigma^{r2}\int_0^T \hat{B}(s,T)\exp(-\Lambda_{sT})ds, \tag{3.16}$$

where $f(0,T)$ is the forward rate at time 0 for borrowing at time T for infinitesimally small amount of time defined as

$$f(0,T) = -\frac{\partial P(0,T)}{\partial T}.$$

Thus, remembering that $r_t = y_t + \bar{y}_t$, using Equation 3.16 it can be seen that

$$r_t = f(0,t) + x_t + \sigma^{r2}\int_0^t \hat{B}(s,t)\exp(-\Lambda_{st})ds. \tag{3.17}$$

Therefore, r_t is a function of x_t, $f(0,t)$, k, and σ^r only. The forward rate is extracted from the yield curve and parameters k and σ^r are obtained from calibration in the Appendix. Thus, we do not need $\bar{x}_t$ or θ_t for the remainder of the work.

3.5.2 Joint density PDE

Local volatility in the SR-LV model setting is given in Equation 3.1. To evaluate this for all K and t, we need to numerically evaluate this expectation using the joint density of interest rate and stock:

$$E^{Q_T}\left[r_T \mathbf{1}_{S_T>K}\right].$$

Consider the following system of SDEs where $x_t = \ln(S_t/F_t)$, thus removing dividends from the problem:

$$\begin{aligned} dx_t &= \left(g_t(y_t) - \frac{1}{2}\sigma^{s2}\right)dt + \sigma^s dW_t^s, \\ dy_t &= -ky_t dt + \sigma^r dW_t^r, \end{aligned}$$

where $g_t(y_t) = r_t - f(0,t)$ (see Equation 3.16), correlation between the Brownian increments is ρ.

In Section 3.3, the PDE for the joint density was derived in a multi-dimensional setting. Because the initial condition for the density is a delta function, it is preferable to work under a transformation of variables and build a PDE for $\phi'(x', y', t)$ where the initial density will be approximately normal.

For the O-U process on the short rate side, the distribution is Gaussian for all

t and using results on Hull-White from the Appendix, the variance of y_t is given by

$$V[y_t] = \sigma^{r2} \frac{1 - e^{-2kt}}{2k}.$$

So, let

$$b_t = \sigma^r \sqrt{\frac{1 - e^{-2kt}}{2k}}.$$

On the equity side, the variance at t is unknown exactly since the marginal distribution will have a skew because of the local volatilities. Thus, following Overhaus *et al.* [23], let

$$a_t = \sigma^s_{atm}(t)\sqrt{t}.$$

where $\sigma^s_{atm}(t)$ is the at-the-money implied volatility at time t. A consequence of using this approximation is that the PDE grid needs to be wider in terms of x'_t to account for the fact that using a_t as a standard deviation alone will not pick up the skew.

The PDE for $\phi'(x', y', t)$ using Equation 3.6 can now be written as (omitting some subscripts):

$$0 = \frac{\partial \phi'}{\partial t} + g_t(by')\phi' + \frac{g_t(by')}{a}\frac{\partial \phi'}{\partial x'} - \frac{1}{2a}\frac{\partial(\sigma^{s2}\phi')}{\partial x'} - \frac{\dot{a}}{a}\frac{\partial(x'\phi')}{\partial x'} -$$
$$\frac{\sigma^{r2}}{2b}\frac{\partial(y'\phi')}{\partial y'} - \frac{1}{2a^2}\frac{\partial^2(\sigma^{s2}\phi')}{\partial x'^2} - \frac{\sigma^{r2}}{2b^2}\frac{\partial^2\phi'}{\partial y'^2} - \frac{\rho\sigma^r}{ab}\frac{\partial^2(\sigma^s\phi')}{\partial x'\partial y'}.$$

Expanding the partial derivatives by product rule and collecting terms we get

$$0 = \frac{\partial \phi'}{\partial t} + q\phi' + q1\frac{\partial \phi'}{\partial x'} + q2\frac{\partial \phi'}{\partial y'} + q11\frac{\partial^2 \phi'}{\partial x'^2} + q22\frac{\partial^2 \phi'}{\partial y'^2} + q12\frac{\partial^2 \phi'}{\partial x'\partial y'}, \tag{3.18}$$

where

$$q = g_t(by') - \frac{1}{2a}\frac{\partial \sigma^{s2}}{x'} - \frac{\dot{a}}{a} - \frac{\sigma^{r2}}{2b^2} - \frac{1}{2a^2}\frac{\partial^2 \sigma^{s2}}{\partial x'^2}$$

$$
\begin{aligned}
q1 &= \frac{g_t(by')}{a} - \frac{\sigma^{s2}}{2a} - \frac{\dot{a}x'}{a} - \frac{1}{a^2}\frac{\partial \sigma^{s2}}{\partial x'}, \\
q2 &= -\frac{\sigma^{r2}y'}{2b^2} - \frac{\rho\sigma^r}{ab}\frac{\partial \sigma^s}{\partial x}, \\
q11 &= -\frac{\sigma^{s2}}{2a^2}, \\
q22 &= -\frac{\sigma^{r2}}{2b^2}, \\
q12 &= -\frac{\rho\sigma^s\sigma^r}{ab}.
\end{aligned}
$$

Now, the boundary conditions and the initial conditions need to be addressed. As for any probability distribution, the probability is nearly zero if looked far enough away from the mean. If x' and y' have the interpretation of standard deviations of x and y, then one can assume that there is no significant probability outside of 4 or 5 standard deviations (i.e., the range of x' and y' can be set from -4 to 4 or -5 to 5). One must keep in mind that for the stock price, the variance is not know at each time t explicitely, so it is safer to define a wider grid. As time increases, the tail (skew) in the marginal distribution lengthens, so the range of x' and y' is taken to be -5 to 5. The PDE is solved on the uniform grid and the boundary value of $\phi'(x', y', t)$ is set to zero.

In a sense, $\phi(x, y)$ has been standardized through the change of variables as in Overhaus *et al.* [23] and therefore, the initial joint density is Gaussian given by:

$$
\begin{aligned}
\phi'(x', y', t) &\approx \frac{ab}{2\pi\sqrt{1-\rho^2}\sigma^s\sigma^r t} \exp\left(-\frac{1}{2(1-\rho^2)} \left(\frac{x'^2a^2}{\sigma^s 2t} + \frac{y'^2b^2}{\sigma^r 2t} - \frac{2\rho x'y'ab}{\sigma^s\sigma^r t} \right) \right) \\
&= \frac{ab}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)} \left(x'^2 + y'^2 - 2\rho x'y' \right) \right)
\end{aligned}
$$

One issue to consider is the time 0. When $t = 0$, some coefficients are undefined, therefore, the PDE solver forward starts at a short time after zero (e.g., 10^{-6}) to avoid this issue and then all coefficient can be evaluated at present time.

Local volatility can be recovered iteratively using:

$$\sigma^2(K,T) = 2\frac{C_T + q(C - KC_K) - KP(0,T)E^{Q_T}\left[r_T \mathbf{1}_{S_T>K}\right]}{K^2 C_{KK}}. \tag{3.19}$$

At time zero, the expectation in Equation 3.19 is known. In implementation, the PDE starts an increment away from zero, but the initial distribution is approximately Gaussian. Thus, local volatility is know at the beginning. Then, all the coefficients are defined and the solver can take a step, recovering the density at the second time period. Local volatility for the second period is then calculated and the solver steps through to find the density at step three. Through such iterative method, we simultaneously recover the joint density and local volatility surface at all times t.

3.5.3 Numerical Results

As an example, consider the T-forward joint density generated for three years in Figure 3.1 in terms of original variables x and y for S&P 500 on June 19, 2002. Effect of positive and negative correlation can be seen in Figures 3.2 and 3.3 respectively.

Rotating this figure, observe that in the interest rate marginals are Gaussian in distribution (Figure 3.5). Stock price marginals have a skew because of the leverage effect in local volatility structure as seen in Figure 3.4. The skew becomes more prominent with time.

This last figure illustrates the local volatility surface extracted by simultane-

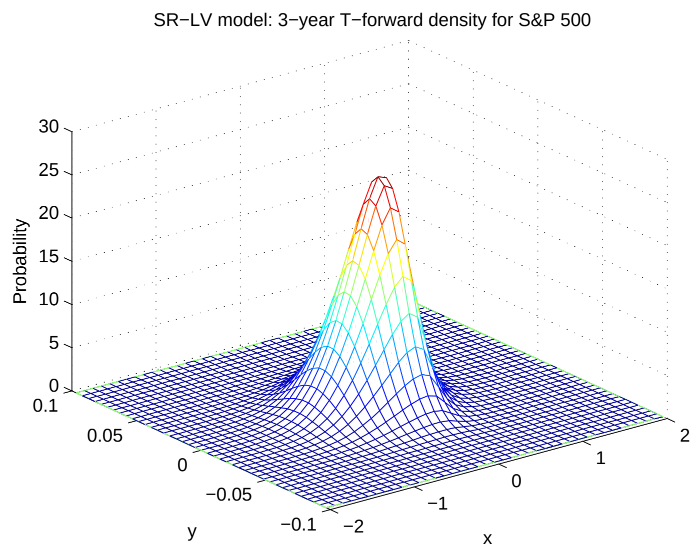

Figure 3.1: SR-LV model: 3-year joint density of x and y for S&P 500 on June 19, 2002 with zero correlation assumption.

ously solving the PDE for the joint density and local volatility. The shape of the surface is consistent with those observed in literature (e.g., Derman [8]).

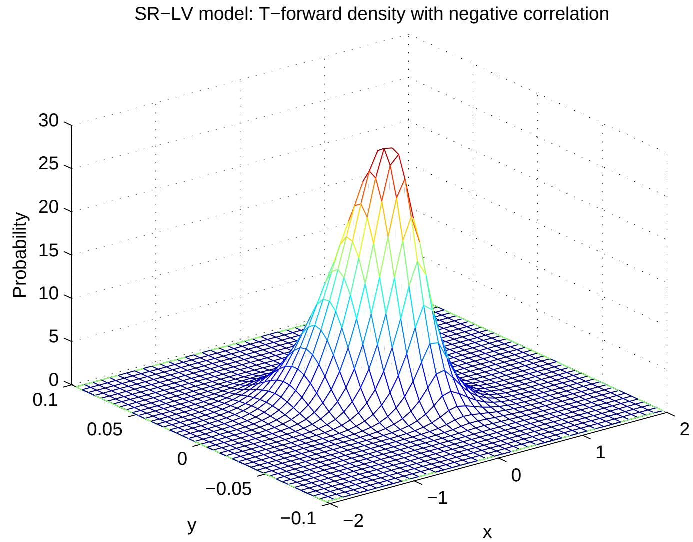

Figure 3.2: SR-LV model: 3-year joint density of x and y for S&P 500 on June 19, 2002 with -.5 correlation assumption.

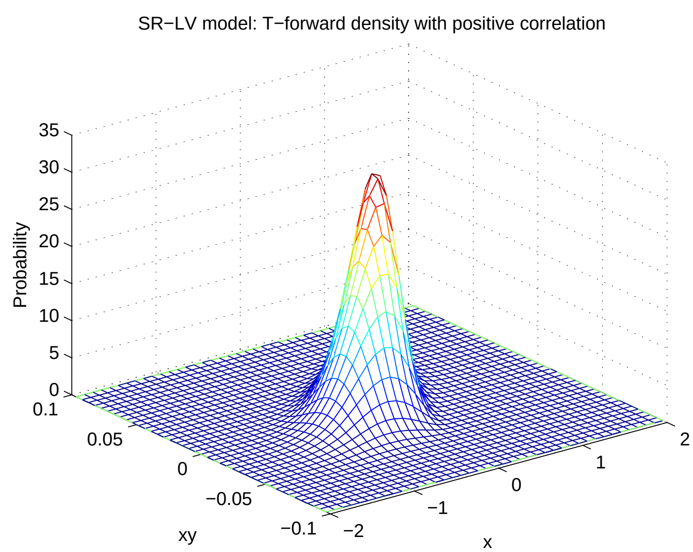

Figure 3.3: SR-LV model: 3-year joint density of x and y for S&P 500 on June 19, 2002 with .5 correlation assumption.

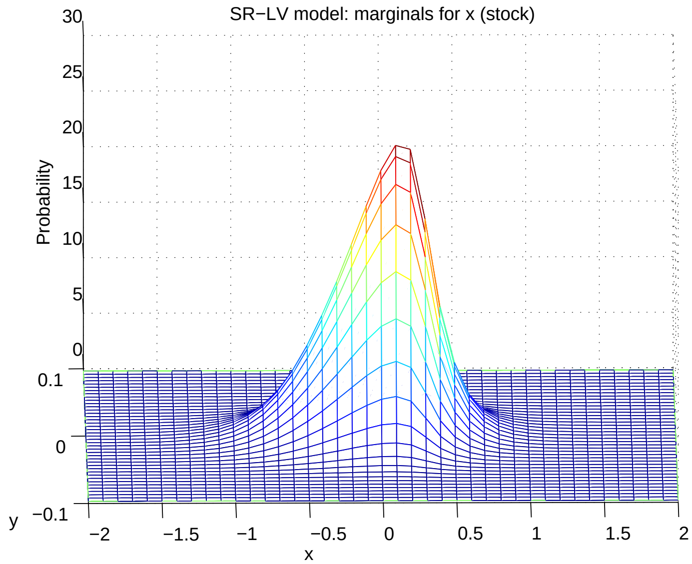

Figure 3.4: SR-LV model: View of marginal densities of x S&P 500 on June 19, 2002 with zero correlation assumption.

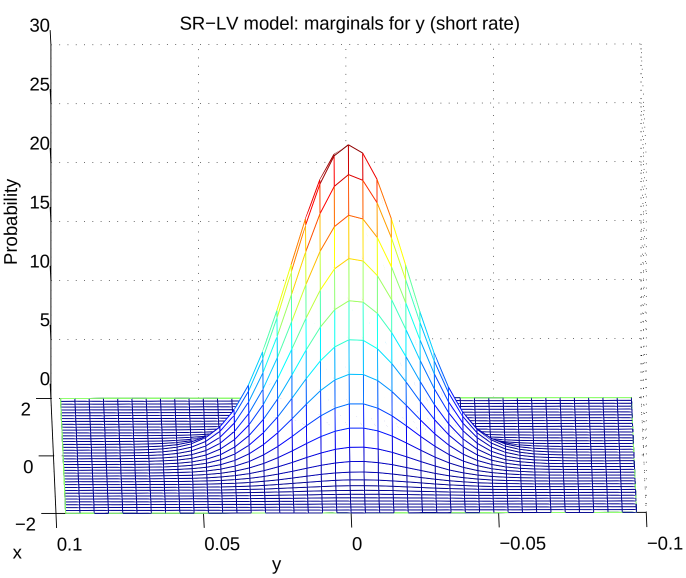

Figure 3.5: SR-LV model: View of marginal densities of y S&P 500 on June 19, 2002 with zero correlation assumption.

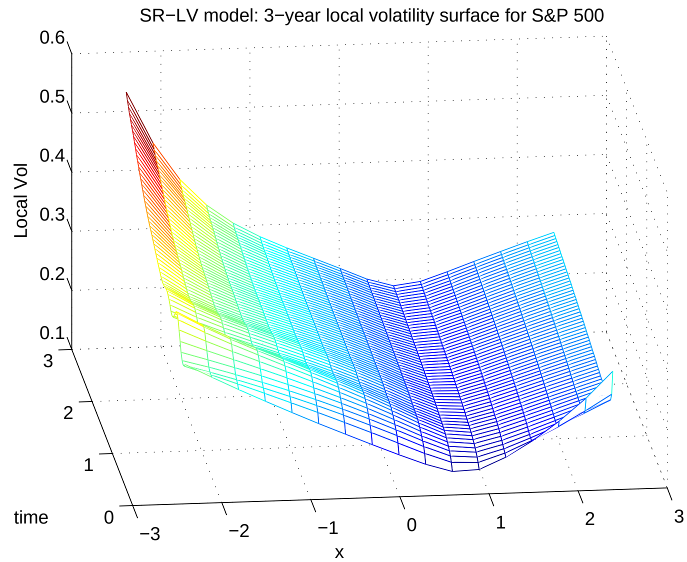

Figure 3.6: SR-LV model: local volatility surface generated in presence of stochastic interest rates for up to 3 years. Zero correlation between stock and interest rate is assumed.

3.6 SR-SLV model

Sometimes it is observed in the market that volatility for some stock changes and the stock price remains unchanged. This effect is not captured in SR-SLV model. One remedy is the introduction of stochastic volatility. But there are two difficulties. First, stochastic volatility models do not capture the implied volatility surface as well and the local volatility model. Second, introducing stochastic volatility means introducing another factor that is correlated with the stock price movements. The consequence for the PDE construction is inclusion of quite a few new terms in three dimensions because of non-zero correlation.

A practical solution to the described phenomenon is the inclusion of another random factor which is uncorrelated with the stock price and interest rates. This random variable follows a log-normal process reverting to some mean, since we do not wish this random shock to be negative. Thus, the SR-SLV model is given by the following system of SDEs:

$$\frac{dS_t}{S_t} = (r_t - q)dt + \sigma_*^s(S,t)Z_t dW_t^s, \tag{3.20}$$

$$dr_t = (\theta_t - kr_t)dt + \sigma^r dW_t^r, \tag{3.21}$$

$$d\ln Z_t = \nu(\zeta_t - \ln Z_t)dt + \lambda dW_t^z, \tag{3.22}$$

$$dW_t^s dW_t^r = \rho dt, \tag{3.23}$$

where dW_t^z is uncorrelated with dW_t^s and dW_t^r. Z_0 is assumed to be 1 and $\sigma_*^s(S,t) = \sigma_*^s$ is the leverage function whose calibration will be discussed later.

3.6.1 Applying Gyöngy

At this time, Gyöngy's [14] result becomes very useful in defining the SR-SLV model dynamics. It is desired to have the same marginal distributions for the stock price for process in Equation 3.20 as those given by the system for the SR-LV model (Equations 3.11, 3.12, and 3.13). Gyöngy gives the tools for such a construction. That is, applying Gyöngy to Equation 3.20 we must have

$$\frac{dS}{S} = (E[r_t|S_t = S] - q)dt + \sqrt{\sigma_*^{s2} E[Z_t^2|S_t = S]}dW_t^s. \tag{3.24}$$

Or, the leverage function must be calibrated so that for all t

$$\sigma^{s2}(K,t) = \sigma_*^{s2}(K,t)E[Z_t^2|S_t = K]. \tag{3.25}$$

Similar to Madan [19], if $\sigma_*^{s2}(K,t)$ has the interpretation of being the average local volatility at time t, then we must force $E[Z_t^2] = 1$ for all t. This will also uncover the value for parameter ζ_t.

Now, Z_t follows a log-normal mean-reverting model. This is just a slightly different form of Hull-White process and Z_t is said to follow the Black-Karasinski model [2]. In this setting, Z_t can only take on positive values. Then, $\ln Z_t$ is normally distributed with

$$\mu^{\ln Z_t} = E[\ln Z_t] = \ln Z_0 e^{-\nu t} + \int_0^t \nu \zeta_t e^{\nu(u-t)} du$$

$$(\sigma^{\ln Z_t})^2 = V[\ln Z_t] = \lambda^2 \left(\frac{1 - e^{-2\nu t}}{2\nu} \right).$$

This implies that Z_t is distributed log-normally with

$$\mu^{Z_t} = E[Z_t] = e^{\mu + (\sigma^{\ln Z_t})^2/2}$$

$$(\sigma^{Z_t})^2 = V[Z_t] = \left(e^{(\sigma^{\ln Z_t})^2} - 1 \right) e^{2\mu + (\sigma^{\ln Z_t})^2},$$

and $E[Z_t^2] = V[Z_t] + E[Z_t]^2$. To have $V[Z_t] + E[Z_t]^2 = 1$ becomes a problem of $\mu^{\ln Z_t} = -(\sigma^{\ln Z_t})^2$ or

$$\int_0^t \nu \zeta_t e^{\nu(u-t)} du = -\frac{\lambda^2}{2\nu} \left(1 - e^{-2\nu t} \right), \tag{3.26}$$

where an Z_0 is assumed to be 1. Multiplying both sides of Equation 3.26 by $e^{\nu t}/\nu$ we obtain

$$\int_0^t \zeta_t e^{\nu u} du = -\frac{\lambda^2}{2\nu^2} \left(e^{\nu t} - e^{-\nu t} \right). \tag{3.27}$$

The choice for ζ_t which satisfies Equation 3.27 is

$$\zeta_t = -\frac{\lambda^2}{2\nu} \left(1 + e^{-2\nu t} \right).$$

A similar approach for finding ζ_t was done in Madan [19].

3.6.2 Joint density PDE

In the SR-SLV setting, local volatility is an input that was computed in the SR-LV model. Then, for a given set of parameters ν and λ in process (3.20), the leverage function, σ_*^s, must be calibrated to the local volatility surface. Because of stochastic interest rates, simulation is done under the T-forward measure, thus using Gyöngy, calibration will be done using this measure in:

$$E^{Q_T}[Z_t^2|S_t = K].$$

This is true, because under a measure change, volatility remains uneffected (via Girsanov's theorem). To evaluate this expectation, we again need to know the joint density, but this time in three variables: stock price, interest rate, and Z.

Consider the following system of SDEs where $x_t = \ln(S_t/F_t)$ and $z_t = \ln Z_t$:

$$\begin{aligned} dx_t &= \Big(g_t(y_t) - \frac{1}{2}(\sigma_*^s Z_t)^2\Big)dt + \sigma_*^s Z_t dW_t^s, \\ dy_t &= -ky_t dt + \sigma^r dW_t^r, \\ dz_t &= \nu(\zeta_t - z_t)dt + \lambda dW_t^z, \end{aligned}$$

where $g_t(y_t) = r_t - f(0,t)$ and the right hand side of $g_t(y_t)$ can be seen from equation (3.17). Correlation between dW^s and dW^r is ρ and dW^z is uncorrelated with the other two Brownian motions.

As in section (3.3), we work under a transformation of variables and build a PDE for $\phi'(x', y', z', t)$ where the initial condition will be approximately normal. For the z_t process, the distribution is Gaussian for all t and using results on Hull-White from the Appendix, the variance of z_t is given by

$$V[z_t] = \lambda^2 \frac{1 - e^{-2\nu t}}{2\nu}.$$

So, let

$$c_t = \lambda \sqrt{\frac{1 - e^{-2\nu t}}{2\nu}}.$$

For the processes x_t and y_t, define a_t and b_t in the same manner as in the SR-LV setting:

$$a_t = \sigma^s_{atm}(t)\sqrt{t},$$
$$b_t = \sigma^r \sqrt{\frac{1 - e^{-2kt}}{2k}},$$

where $\sigma^s_{atm}(t)$ is the at-the-money implied volatility at time t. Using this simplification is amplified in the presence of Z_t, since more skewness is present for higher levels of z_t.

The PDE for $\phi'(x', y', z', t)$ using equation (3.6) can now be written as (omitting some subscripts):

$$0 = \frac{\partial \phi'}{\partial t} + (g_t(by') - \nu)\phi' + \frac{g_t(by')}{a}\frac{\partial \phi'}{\partial x'} - \frac{1}{2a}\frac{\partial(\sigma^{s2}_* Z^2 \phi')}{\partial x'} - \frac{\dot{a}}{a}\frac{\partial(x'\phi')}{\partial x'} - \frac{\sigma^{r2}}{2b}\frac{\partial(y'\phi')}{\partial y'}$$
$$\left(\frac{\nu\zeta}{c} - kz'\right)\frac{\partial \phi'}{\partial z'} - \frac{1}{2a^2}\frac{\partial^2(\sigma^{s2}_* Z^2 \phi')}{\partial x'^2} - \frac{\sigma^{r2}}{2b^2}\frac{\partial^2 \phi'}{\partial y'^2} - \frac{\lambda^2}{2c^2}\frac{\partial^2 \phi'}{\partial z'^2} - \frac{\rho\sigma^r}{ab}\frac{\partial^2(\sigma^s_* Z\phi')}{\partial x'\partial y'}.$$

Expanding the partial derivatives by product rule and collecting terms we get

$$0 = \frac{\partial \phi'}{\partial t} + q\phi' + q1\frac{\partial \phi'}{\partial x'} + q2\frac{\partial \phi'}{\partial y'} + q3\frac{\partial \phi'}{\partial z'} + q11\frac{\partial^2 \phi'}{\partial x'^2} + q22\frac{\partial^2 \phi'}{\partial y'^2} + q33\frac{\partial^2 \phi'}{\partial z'^2} + q12\frac{\partial^2 \phi'}{\partial x'\partial y'}, \tag{3.28}$$

where

$$q = g_t(by') - \frac{e^{2z'c}}{2a}\frac{\partial \sigma^{s2}}{x'} - \frac{\dot{a}}{a} - \frac{\sigma^{r2}}{2b^2} - \frac{e^{2z'c}}{2a^2}\frac{\partial^2 \sigma^{s2}}{\partial x'^2} - kz',$$

$$q1 = \frac{g_t(by')}{a} - \frac{e^{2z'c}\sigma^{s2}}{2a} - \frac{\dot{a}x'}{a} - \frac{e^{2z'c}}{a^2}\frac{\partial\sigma^{s2}}{x'},$$
$$q2 = -\frac{\sigma^{r2}y'}{2b^2} - \frac{\rho\sigma^r e^{z'c}}{ab}\frac{\partial\sigma^s}{\partial x},$$
$$q3 = \frac{\nu\zeta}{c} - kz',$$
$$q11 = -\frac{e^{2z'c}\sigma^{s2}}{2a^2},$$
$$q22 = -\frac{\sigma^{r2}}{2b^2},$$
$$q33 = -\frac{\lambda^2}{2c^2},$$
$$q12 = -\frac{\rho e^{z'c}\sigma^s\sigma^r}{ab}.$$

The boundary conditions and the initial conditions addressed similarly as in the SR-LV setting. However, the presence of Z admits larger skew in the marginal distribution of the stock. Therefore, if x', y', and z' intuitively represent the standard deviations away from the mean for the respective distributions, then it was observed that the range can be set at -4 to 4 (or -5 to 5 for longer maturities) for each variable to pick up all significant probabilities. The PDE is solved on the uniform grid and the boundary value of $\phi'(x', y', z', t)$ is set to zero. Note, that it is not a necessity to work on a uniform grid, since the defined range is too big for y' and z'. The large range is due to the skew in the stock and no great efficiency is lost by working on a uniform grid.

Following the same logic as for SR-LV model the initial joint density is Gaussian for small t given by:

$$\phi'(x', y', z', t) \approx \frac{ab}{(2\pi)^{\frac{3}{2}}\sqrt{1-\rho^2}\sigma^s\sigma^r\lambda t}\exp\Bigg(-\frac{1}{2(1-\rho^2)}\left(\frac{x'^2a^2}{\sigma^s 2t} + \frac{y'^2b^2}{\sigma^r 2t} - \frac{2\rho x'y'ab}{\sigma^s\sigma^r t}\right) - \frac{1}{2}\frac{z'^2c^2}{\lambda^2 t}\Bigg) = \frac{ab}{(2\pi)^{\frac{3}{2}}\sqrt{1-\rho^2}}\exp\Bigg(-\frac{1}{2(1-\rho^2)}\left(x'^2 + y'^2 - 2\rho x'y'\right) - \frac{z'^2}{2}\Bigg).$$

Again, for $t = 0$, some coefficients are undefined. Therefore, the PDE solver starts at a short time after zero (e.g., 10^{-6}) to avoid discontinuities and then all coefficient can be evaluated at present time.

Using local volatility from SR-LV model, the leverage function σ_*^2 can be iteratively recovered by stepping through the solver and evaluating the expectation $E^{Q_T}[Z_t^2|S_t = K]$. For example, at time 0 (or 10^{-6}) the expectation is known via initial distribution and the leverage function is recovered. Then using parameter values at time, the solver obtains the solution at the second step, where the leverage function can again be recovered.

3.6.3 Numerical Results

It is a bit difficult to visualize the density in three variables, so Figure 3.7 just illustrates the 3-year conditional joint density of x and y for z=0. Parameters for the z_t process are $\nu = .5$ and $\lambda = .5$. Figures 3.8 and 3.9 show the conditional joint density of x and y for large (2 standard deviation to the right of the mean) and small (2 standard deviation to the left of the mean) z respectively.

First, the marginal distributions in y and z are Gaussian. This feature is dictated by their respective processes with constant volatility. Second, there is still a skew to the left in the marginal distributions of x.

However, the presence of the random factor of volatility z gives rise to another interesting characteristic. When z is large, the conditional marginal distributions of x begin to take a bimodal shape. The bimodal nature is not uncommon and is well documented in literature (e.g., Derman [8], Jackwerth [24]). Li and Pearson [18] remark that the bimodal feature of the density is a necessary stylistic characteristic for models involving stochastic volatility. For small z, the conditional marginal distributions of x and y are very thin reflecting the low level of volatility.

In the SR-SLV model, the calibrated leverage function picks up the leverage effect (i.e., inverse relationship between volatility and stock price movements). The random component, z, is a random shock to volatility that is uncorrelated with the movement in the stock price. In the context of the S&P 500, large z reflects an upward spike in volatility. With an upward spike in volatility, the stock price will move and the leverage effect dictates a likely downward move. But, with a large sudden move in volatility, the market is likely to stabilize at some supported lower level. Thus, the bimodal feature of the distribution of the index can intuitively be explained by this market behavior. Some market participants believe that the distribution of the S&P 500 may even be trimodal.

While it is comforting that some market behavior is reflected in the SR-SLV model, the bimodal feature really arises from the nature of model. When z is large, the variance is big. Variance itself involves a square of the price change and when this number is big there could be either a large positive or negative move in the price.

As expected, when the volatility of volatility, λ is decreased, the bimodal feature disappears. This happens because the SR-SLV model loses it's stochastic nature and converges to the SR-LV model as $\lambda \to 0$. That is, the calibrated leverage function converges to the local volatility surface from SR-LV model.

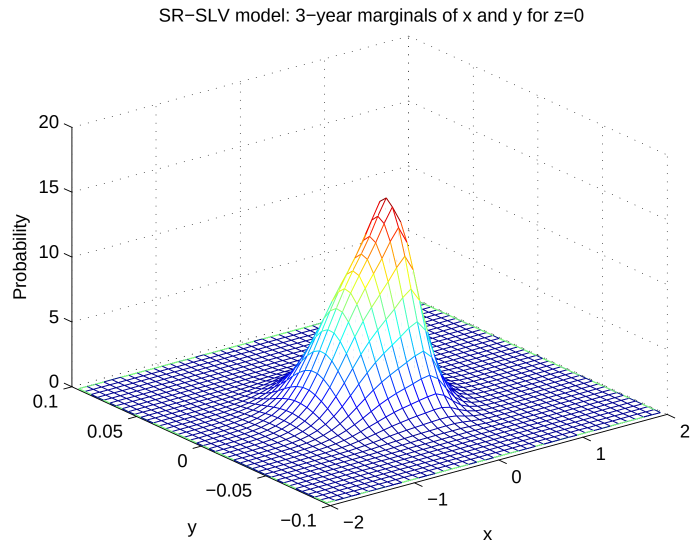

Figure 3.7: SR-SLV model: View of marginal density of x and y when z=0. Note that these is still the skew present but through time, the density is developing a second mode. Zero correlation between stock and interest rate is assumed.

The plots of the conditional marginal densities are for the calibrated parameters of the short rate. However, as the short rate volatility increases, the bimodal feature become more pronounced and can be seen for earlier maturities. This analysis is postponed until the discussion of realized variance options across various short rate volatilities in Chapter 4.

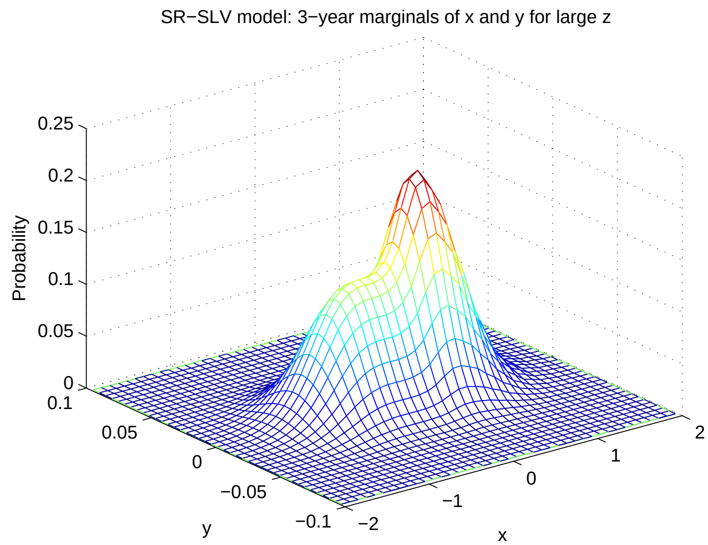

Figure 3.8: SR-SLV model: View of marginal density of x and y when z is large. For larger shocks in z, the distribution is more bimodal. Zero correlation between stock and interest rate is assumed.

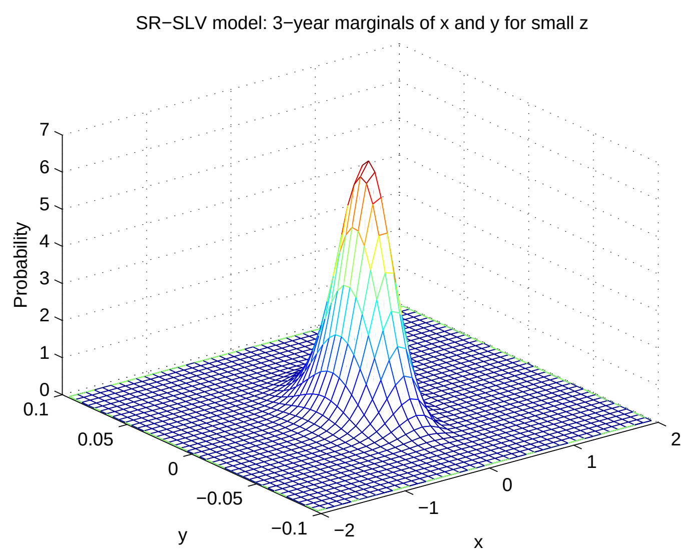

Figure 3.9: SR-SLV model: View of marginal density of x and y when z is small. For small shocks in z, the distribution is thiner when volatility is small. Zero correlation between stock and interest rate is assumed.

Chapter 4

Pricing Variance/Volatility Derivatives

4.1 Risk-neutral to T-forward measure change

The pricing of variance derivatives is done under the T-forward measure so that we could utilize the discount bond as a numeraire and avoid the cumbersome calculations involving the money market account. This section will invoke the Girsanov's theorem, where the goal will be to identify the form of the Radon-Nikodym derivative involving the previsible process necessary for the measure change.

First, let us address the measure change for the Hull-White interest rate process, which can be written as an O-U process as in Chapter 3:

$$dy_t = -ky_t dt + \sigma^r dW_t^r, \tag{4.1}$$

where now

$$r_t = y_t + \bar{y}_t,$$

and

$$d\bar{y}_t = (\theta_t - k\bar{y}_t)dt.$$

The Radon-Nikodym derivative for changing from risk-neutral to T-forward measure is just the ratio of the numeraires, money market account and discount bond, respectively. Overhaus *et al.* shows how to derive the Radon-Nikodym derivative in this setting using result from the Hull-White section in Chapter 3. Starting with the money market account:

$$
\begin{aligned}
B_T &= \exp\Big(\int_0^T r_t dt\Big) \\
&= \exp\Big(\int_0^T (y_t + \bar{y}_t)dt\Big) = \frac{1}{P(0,T)} \exp\Big(\int_0^T y_t dt + \frac{1}{2}\int_0^T \sigma^{r2}\hat{B}(t,T)^2 dt\Big) \\
&= \frac{1}{P(0,T)} \exp\Big(\int_0^T \int_0^t \sigma^r \exp(-\Lambda_{us}) dW_u dt + \frac{1}{2}\int_0^T \sigma^{r2}\hat{B}(t,T)^2 dt\Big) \\
&= \frac{1}{P(0,T)} \exp\Big(\int_0^T \sigma^r \hat{B}(t,T) dW_t + \frac{1}{2}\int_0^T \sigma^{r2}\hat{B}(t,T)^2 dt\Big)
\end{aligned}
$$

where $\Lambda_{ts} = k(s-t)$ and $\hat{B}(t,T) = \frac{1-e^{k(t-T)}}{k}$. Substitutions in the above derivation are made from the work in Section 3.5.1 deriving the expression for the discount bond price.

The Radon-Nikodym derivative is

$$
\frac{dQ_T}{dQ} = \frac{1}{B_T P(0,T)} = \exp\Big(-\int_0^T \sigma_t \hat{B}(t,T) dW_t - \frac{1}{2}\int_0^T \sigma_t^2 \hat{B}(t,T)^2 dt\Big) \ .
$$

The previsible process for the measure change is then $\gamma_t^r = \sigma_t^r \hat{B}(t,T)$. Thus, under the T-forward measure

$$
W_t^{rQ_T} = W_t^r + \int_0^T \gamma_t^r dt.
$$

Under the T-forward measure, equation (4.1) becomes

$$
dy_t = -\Big(ky_t + \sigma^{r2}\hat{B}(t,T)\Big)dt + \sigma^r dW_t^{rQ_T}. \tag{4.2}
$$

Now, consider the setting with the equity process and stochastic interest rates:

$$
\begin{aligned}
&\frac{dS_t}{S_t} = (r_t - q)dt + \sigma^s dW_t^s, \\
&dy_t = -ky_t dt + \sigma^r dW_t^r, \\
&dW_t^s dW_t^r = \rho dt.
\end{aligned}
$$

To simplify this to a problem without dividends, let $X_t = S_t\ e^{qt}$. Then our system has the following dynamics:

$$\frac{dX_t}{X_t} = r_t dt + \sigma_t^s dW_t^s,$$
$$\frac{dP(t,T)}{P(t,T)} = r_t dt - \sigma^r \hat{B}(t,T) dW_t^r,$$
$$dW_t^s dW_t^r = \rho dt,$$

where $P(t,T)$ is the price of a discount bond at time t, maturing at T.

At this time, Itô's quotient rule can be utilized to determine the change in the drift under the change in measure. Under T-forward measure, the forward price $X_t/P(t,T)$ will be a martingale and we can use Ito's quotient rule to get γ_t^s for the equity measure change. Itô's quotient rule is given by:

$$\frac{d(X_t/P(t,T))}{X_t/P(t,T)} = \frac{dX_t}{X_t} - \frac{dP(t,T)}{P(t,T)} + \frac{dP(t,T)dP(t,T)}{P(t,T)^2} - \frac{dX_t dP(t,T)}{X_t P(t,T)}$$
$$= (\mu_x - \mu_p + \sigma_p^2 - \rho\sigma_x\sigma_p)dt + \sigma_x dW_t^r - \sigma_p dW_t^S\ .$$

In this particular setting:

$$\frac{d(X_t/P(t,T))}{X_t/P(t,T)} = \Big(r_t - r_t + \sigma_t^{r2}\hat{B}(t,T)^2 - \rho\sigma_t^r\sigma_t^s\hat{B}(t,T)\Big)dt + \sigma_t^s dW_t^s + \sigma_t^r\hat{B}(t,T)dW_t^r\ .$$

We need to have zero drift and we know from earlier that $\gamma_t^r = \sigma_t^r\hat{B}(t,T)$. It follows that $\gamma_t^s = -\rho\sigma_t^r\hat{B}(t,T)$. The resulting processes undet T-forward measure are:

$$\frac{dX_t}{X_t} = (r_t + \rho\sigma^r\sigma_t^s\hat{B}(t,T))dt + \sigma_t^s dW_t^{sQ_T},$$
$$\frac{dP(t,T)}{P(t,T)} = (r_t + \sigma^{r2}\hat{B}(t,T)^2)dt - \sigma^r\hat{B}(t,T)dW_t^{rQ_T}.$$

In the SR-SLV model, there is one more factor Z_t, the random component of stochastic local volatility. $Z_t = \exp(z_t)$, where

$$dz_t = \nu(\zeta_t - z_t)dt + \lambda dW_t^z.$$

But, z_t is not a function of r_t and it is uncorrelated with the other processes. Therefore, from the Itô's quotient it can be seen that under the T-forward measure we have

$$dz_t = \nu(\zeta_t - z_t)dt + \lambda dW_t^{zQ_T}.$$

In summary, for the SR-LV, the processes under the T-forward measure have the form:

$$\begin{aligned}
dx_t &= \left(g_t(y_t) - \frac{1}{2}\sigma_t^{s2} + \rho\sigma^r\sigma_t^s\hat{B}(t,T)\right)dt + \sigma_t^s dW_t^{sQ_T}, \\
dy_t &= -\left(ky_t + \sigma^{r2}\hat{B}(t,T)\right)dt + \sigma^r dW_t^{rQ_T}, \\
dW_t^{sQ_T} dW_t^{rQ_T} &= \rho dt,
\end{aligned}$$

where σ_t^s represent the local volatility and $g_t(y_t) = r_t - f(0,t)$. Similarly, for the SR-SLV model, changing measure yields:

$$\begin{aligned}
dx_t &= \left(g_t(y_t) - \frac{1}{2}\sigma_{*t}^{s2}e^{2z_t} + \rho\sigma^r\sigma_{*t}^s e^{z_t}\hat{B}(t,T)\right)dt + \sigma_{*t}^s e^{z_t} dW_t^{sQ_T}, \\
dy_t &= -\left(ky_t + \sigma^{r2}\hat{B}(t,T)\right)dt + \sigma^r dW_t^{rQ_T}, \\
dz_t &= \nu(\zeta_t - z_t)dt + \lambda dW_t^{zQ_T}, \\
dW_t^{sQ_T} dW_t^{rQ_T} &= \rho dt,
\end{aligned}$$

where dW^{zQ_T} is uncorrelated with the other Brownian motions.

4.2 Intuition behind variance swaps and correlation

The two models considered in this dissertation will actually give the same price for variance swaps (shown in Section 4.4). Thus, if is sufficient to just analyze the SR-LV model for the moment. This section will examine the price for variance swaps for different levels of correlation and across time with varying volatility of the stochastic interest rate.

Consider the processes in SR-LV model under the T-forward measure:

$$\frac{dS_t}{S_t} = \Big(r_t - q + \rho\sigma^s_{LV}\sigma^r\hat{B}(t,T)\Big)dt + \sigma^s_t dW^{sQ_T}_t, \tag{4.3}$$

$$dy_t = -ky_t dt - \sigma^{r2}\hat{B}(t,T)dt + \sigma^r dW^{rQ_T}_t, \tag{4.4}$$

where

$$r_t = y_t + f(0,t) + \int_0^t \sigma^{r2}\hat{B}(s,t)\exp(-\Lambda_{st})ds, \tag{4.5}$$

$$dW^{sQ_T}_t dW^{rQ_T}_t = \rho dt,$$

and $\hat{B}(s,t)$, Λ_{st} are defined the same as in the previous section and σ^s_t is the local volatility. Applying Gyöngy [14] to the process in 4.3 results in

$$\frac{dS_t}{S_t} = \Big(E[r_t|S_t] - q + \rho\sigma^s_t\sigma^r\hat{B}(t,T)\Big)dt + \sigma^s_t dW^{sQ_T}_t. \tag{4.6}$$

Proposition: Under the T-forward measure, the expected short rate, r_t, is equal to the forward rate, $f(0,t)$.

Proof: To see this take expectation of both sides of equation 4.5 and we get:

$$E[r_t] = E[y_t] + f(0,t) + \int_0^t \sigma^{r2}\hat{B}(s,t)\exp(-\Lambda_{st})ds.$$

Now, the solution to equation (4.4), with $y_0 = 0$, is given by the following (details of distribution under Hull-White assumptions are given in the Appendix):

$$y_t = -\int_0^t \sigma^{r2}\hat{B}(s,t)\exp(-\Lambda_{st})ds + \int_0^t \sigma^r\exp(-\Lambda_{st})dW^{rQ_T}_s.$$

Thus,

$$E[y_t] = \int_0^t \sigma^{r2}\hat{B}(s,t)\exp(-\Lambda_{st})ds,$$

and $E[r_t] = f(0,t)$.

For simplicity, consider the case of $\rho = 0$. The stock price and the short rate are uncorrelated, but not independent since the return distribution of the stock is non-Gaussian. Thus we cannot conclude that $E[r_t|S_t] = E[r_t] = f(0,t)$. If this were true, we would expect the variance swap price to be the same as under pricing using the log-contract hedge. However, if in the the Hull-White process, the speed of mean reversion (k) is high and the variance (σ^r) is low, then the variance swap price should approach the price using the log-contract hedge from below. If in the Hull-White process we increase the volatility and/or lower the speed of mean reversion, we would expect the difference between SR-LV model price and price via the log-contract hedge for variance swap to grow. Thus, the first observation is that the difference in pricing a variance swap via SR-LV model and log contract depends on "stochasticity" of the short rate process.

Suppose that $\rho \neq 0$. If $\rho > 0$, then the stock return process now has a positive contribution to the drift in $\rho\sigma_t^s\sigma^r\hat{B}(t,T)$. Furthermore, as the stock price increases, the short rate will tend to increase, again positively contributing to the drift. Therefore, under positive correlation the stock will drift higher attaining lower variances. And, when the variances are averaged, we will obtain a lower realized variance that under the $\rho = 0$ assumptions.

The case of $\rho < 0$ is the opposite of the above explanation. In this scenario, there is now a negative contribution to the drift, $\rho\sigma_t^s\sigma^r\hat{B}(t,T)$, and as the stock price rises, the short rate will tend to fall, not adding as much to the drift. Lower stock prices attained should lead to a higher price for the variance swap due to the leverage effect. That is, volatility increases as the stock price falls.

An argument could be made that when the the stock price falls, the short rate will tend to fall when $\rho > 0$ and rise when $\rho < 0$, thus partially nullifying the above argument. However, when the stock price falls, the main driver of the process in Equation 4.3 becomes the volatility due to the leverage effect. In summary, there is an inverse relationship between the variance swap price and correlation: as correlation moves in the positive direction, variance swap price will decrease.

Intuitively, the term structure of variance swaps will deviate from the prices given by the log-contract hedge depending on the "stochasticity" of the short rate. That is, as k increases and σ^r decreases the SR-LV model reduces to deterministic short rates, where under zero correlation assumption, the short rate is equal to the forward rate. So, variance swaps price in this setting should be close to the price via the log-contract hedge. As the short rate model becomes more "stochastic" (e.g., volatility increasing), SR-LV model's price deviations will increase from the replicating price using the log contract.

4.3 Simulation for SR-LV model

Simulation is needed to price variance swaps and variance options. The idea is to simulate paths and collect the local variance at each point of a given path. The average of realized variances over many paths is a good approximation for realized variance. Since expectations are taken under the T-forward measure for pricing (see Chapter 1), the paths must be simulated under the T-forward measure. Two simulation approaches are discussed in this section - simple Euler and Milstein. The Euler method is preferred since the Milstein scheme does not significantly improve accuracy and it involves the additional aspect of bounding the derivative of local volatility.

Thus, we could simulate individual processes:

$$dx_t = \left(g_t(y_t) - \frac{1}{2}\sigma_t^{s2} + \rho\sigma^r\sigma_t^s\hat{B}(t,T)\right)dt + \sigma_t^s dW_t^{sQ_T}, \tag{4.7}$$

$$dy_t = -\left(ky_t + \sigma^{r2}\hat{B}(t,T)\right)dt + \sigma^r dW_t^{rQ_T}, \tag{4.8}$$

$$dW_t^{sQ_T} dW_t^{rQ_T} = \rho dt, \tag{4.9}$$

where σ_t^s represents the local volatility and $g_t(y_t) = r_t - f(0,t)$.

However, there is a better way to tackle this problem where only one process needs to be simulated. Gyöngy's work can be utilized along with the T-forward density that was recovered from Chapter 3. The goal is to write down an equity process that has the same marginals as Equations 4.7-4.9. If Gyöngy's result is applied to just equity then we get:

$$dx_t = \left(E^{Q_T}[g_t(y_t)|x_t] - \frac{1}{2}\sigma_t^{s2} + \rho\sigma^r\sigma_t^s\hat{B}(t,T)\right)dt + \sigma_t^s dW_t^{sQ_T}, \tag{4.10}$$

or substituting for $g_t(y_t)$

$$dx_t = \left(E^{Q_T}[r_t|x_t] - f(0,t) - \frac{1}{2}\sigma_t^{s2} + \rho\sigma^r\sigma_t^s\hat{B}(t,T)\right)dt + \sigma_t^s dW_t^{sQ_T}. \tag{4.11}$$

At this stage Equation 4.11 can be discretized using the Euler or the Milstein schemes. The Euler discretization for $0 = t_0 < t_1 < .... < t_m$ with $h = (t_{i+1} - t_i)$ is given by

$$x(t_{i+1}) = x(t_i)+\left(E^{Q_T}[r(t_i)|x(t_i)]-f(0,t_i)-\frac{1}{2}\sigma^2(S(t_i),t_i)+\rho\sigma^r\sigma(S(t_i),t_i)\hat{B}(t_i,T)\right)h$$

$$+\sigma(S(t_i),t_i)\sqrt{h}N_{i+1}, \tag{4.12}$$

where $N_1, N_2...$ are independent standard normal random variables at times t_1, t_2 and so on (Z is typically used for a standard normal random variance, but it will be confusing with the random factor Z later). Under the Milstein scheme, we add one more term

$$x(t_{i+1}) = x_(t_i)+\left(E^{Q_T}[r(t_i)|x(t_i)]-f(0,t_i)-\frac{1}{2}\sigma^2(S(t_i),t_i)+\rho\sigma^r\sigma(S(t_i),t_i)\hat{B}(t_i,T)\right)h$$

$$+\sigma(S(t_i),t_i)\sqrt{h}N_{i+1}+\frac{1}{2}\frac{\partial\sigma_t^s}{\partial x}\sigma(S(t_i),t_i)h(N_{i+1}^2-1).$$

The stock price can then be recovered by:

$$S(t_{i+1}) = S(0)\frac{\exp(x(t_{i+1})-d(t_{i+1}))}{P(0,t_{i+1})}.$$

Both are implemented, however as confirmed in Glasserman [13], the Milstein scheme does not add much more accuracy. The Euler scheme is weak order 1 where the diffusion term is only expanded to $O(\sqrt{h})$. The Milstein sheme expands the diffusion term to $O(h)$ through the extra term and is considered strong order 1 convergent. Glasserman [13] explains that the weak order of convergence for both schemes is 1, thus the Milstein method does not greatly improve accuracy.

Simulations are not expensive in terms of computational time. For example, 100,000 simulations for a 1-year swap takes about 7 seconds on a home PC. Thus, to reduce variance, the number of simulations can be increased without great increase in computing time.

There is one practical consideration in implementing the simulations. The local volatility surface is defined over a wide range of strikes coming from the PDE solver. For example, the surface is defined in SR-LV setting for strikes within 4 standard deviations of the mean for the one year time horizon. Furthermore, local volatility surface is about twice as steep as implied volatility surface as a function of strike. When the Monte Carlo simulates a path, a cubic spline is used to obtain values of local volatility from the surface.

Some paths of the simulation may take the stock price outside of the range where the local volatility is defined. Because of the steepness of the local volatility curve for decreasing strikes, it is not advisable to use the cubic spline to extrapolate local volatility outside the surface. Moreover, even on the local volatility surface, very large volatilities exist very far away from the mean. These are highly unlikely to ever be observed in the market for S&P 500 (e.g., 1,000% volatility). While these high values are necessary for defining the density in the PDE solver, they are unrealistic for the simulations and can direct the simulation of the stock into extremely large territory (also unrealistic).

There are two remedies to the above problem. One is to bound the local volatility explicitly by some number (e.g., 200%). However, this number is a bit arbitrary. It may not be appropriate to use one number as a bound for all times. The approach taken in this work is the following: if at any time t, the stock price in a simulation ventures beyond 2.5 time-t standard deviations from the mean, then chose the local volatility of 2.5 time-t standard deviations boundary. At any given time, the standard deviation for the stock is approximated as $\sigma_{imp}\sqrt{t}$ since a closed form is not know because of the skew. There is nothing special about 2.5 standard deviations chosen, 3 works just as well. Since $\sigma_{imp}\sqrt{t}$ is an approximation, one should err on the side of caution and use a larger number of deviations.

The practical problem discussed is extended in the Milstein method, where there now exists a partial derivative of local volatility with respect to x. This quantity must also be bounded because of the steepness of the local volatility curve. If the stock price ventures outside the area specified above, then the partial derivative is taken to be the one calculated on the boundary. While bounding the local volatility can intuitively and mathematically be justified, it is difficult to come up with a number to bound the first derivative of local volatility. This is an additional reason for using the simple Euler scheme, since no great improvement in accuracy was observed.

4.4 Simulation for SR-SLV model

In SR-SLV model, computation time is an issue. That is, first the SR-LV model is run to extract the local volatility surface and then the leverage function is calibrated to fit this surface. Thus, two PDE solvers must be run. To save computation time, larger steps can be taken in the simulation, but this comes at the expense of increased variance in the simulation results. From the previous subsection, Milstein method does not provide great improvement in terms of accuracy. Glasserman [13] describes a second order method. However, same computational difficulties persist with first and second partial derivatives of the leverage function as did with the derivatives of local volatility in SR-LV setting. Therefore, again Euler method is preferred with increased number of simulations to improve accuracy. For completeness, the second order method is presented.

For the SR-SLV model, there are three processes that need to be simulated:

$$dx_t = \left(g_t(y_t) - \frac{1}{2}\sigma_{*t}^{s2}e^{2z_t} + \rho\sigma^r\sigma_{*t}^s e^{z_t}\hat{B}(t,T)\right)dt + \sigma_{*t}^s e^{z_t} dW_t^{sQ_T}, \quad (4.13)$$

$$dy_t = -\left(ky_t + \sigma^{r2}\hat{B}(t,T)\right)dt + \sigma^r dW_t^{rQ_T}, \quad (4.14)$$

$$dz_t = \nu(\zeta_t - z_t)dt + \lambda dW_t^{zQ_T}, \quad (4.15)$$

$$dW_t^{sQ_T} dW_t^{rQ_T} = \rho dt, \quad (4.16)$$

where dW^{zQ_T} is uncorrelated with the other Brownian motion increments.

Similar to the SR-SLV, Gyöngy's result can be applied to price variance swaps. In this framework, the single process for x_t that has the same marginals as equations (4.16)-(4.18) and let $Z_t = e^{z_t}$

$$dx_t = \left(E^{Q_T}\left[g_t(y_t) - \frac{1}{2}\sigma_{*t}^{s2}Z_t^2 + \rho\sigma^r\sigma_{*t}^s Z_t\hat{B}(t,T)|x_t\right]\right)dt + \sigma_{*t}^s\sqrt{E[Z_t^2|x_t]}dW_t^{sQ_T}, \quad (4.17)$$

or substituting for $g_t(y_t)$

$$dx_t = \left(E^{Q_T}[r_t|x_t] - f(0,t) - \frac{1}{2}\sigma_{*t}^{s2}E^{Q_T}[Z_t^2|x_t] + \rho\sigma^r\sigma_{*t}^s\sqrt{E^{Q_T}[Z_t^2|x_t]}\hat{B}(t,T)\right)dt \quad (4.18)$$

$$+\sigma_{*t}^s\sqrt{E^{Q_T}[Z_t^2|x_t]}dW_t^{sQ_T}.$$

Now, expected realized variance in the SR-LV model is given by:

$$K_{vs}^{SR-LV} = \frac{1}{T}E^{Q_T}\left[\int_0^T \sigma_u^{s2} du\right],$$

and in the SR-LV model

$$K_{vs}^{SR-SLV} = \frac{1}{T} E^{Q_T} \left[\int_0^T \sigma_{*u}^{s2} E^{Q_T} [Z_u^2 | x_u] du \right].$$

However, it is clear that $K_{vs}^{SR-LV} = K_{vs}^{SR-SLV}$, by interchanging integration and expectation, and using law of total expectations, and remembering that ζ_t was fixed so that $E[Z_t^2] = 1$ for all t. Therefore, for pricing variance swaps, the two models will produce the same price. The two models will produce different variance option prices. However, for SR-SLV model, Equation 4.18 cannot be used for simulation for options because its volatility component now lacks vol of vol in this form. Thus, all three processes in Equations 4.13-4.15 need to be simulated.

First, the processes for y_t and and z_t are Gaussian and can be simulated exactly. Then consider the simple Euler scheme for x_t. The three processes can then be discretized in the following manner with $Z(t_i) = \exp(z(t_i))$:

$$x(t_{i+1}) \approx x(t_i) + \left[y(t_i) + \frac{\sigma^{r2}}{k^2} \left(\frac{1}{2}(e^{-2kt_i} + 1) - e^{-kt_i} \right) - \frac{1}{2} \sigma_*^2 (S(t_i), t_i) Z^2(t_i) \right. \tag{4.19}$$

$$\left. + \rho \sigma^r \sigma_* (S(t_i), t_i) Z(t_i) \hat{B}(t, T) \right] h + \sigma_* (S(t_i), t_i) Z_t \sqrt{h} N_{i+1}^s,$$

$$y(t_{i+1}) = y(t_i) e^{-kh} + \mu_r(t_i) + \sqrt{\frac{1 - e^{-2kdt}}{2k}} (\sigma_1 N_{i+1}^s + \sigma_2 N_{i+1}^r), \tag{4.20}$$

$$z(t_{i+1}) = z(t_i) e^{-\nu h} + \mu_z(t_i) + \sqrt{\frac{1 - e^{-2\nu dt}}{2\nu}} \lambda N_{i+1}^z, \tag{4.21}$$

where N^s,N^r,N^z are uncorrelated standard normal random variables, $\sigma^r = \sigma_1^2 + \sigma_2^2$, and $\sigma_1 = \rho \sigma^r$. Also,

$$\mu_r(t_i) = -\frac{\sigma^{r2}}{k^2} \left(\frac{1}{2}(e^{-2kdt} + 1) - e^{-kdt} \right),$$

$$\mu_z(t_i) = -\frac{\lambda^2}{2\nu} \left(e^{-2\nu t_i} + e^{-\nu h} - e^{-\nu(t_i + t_{i+1})} - 1 \right),$$

and $h = (t_{i+1} - t_i)$. Here, we have used the fact that

$$g_t(y_t) = r_t - f(0,t) = y_t + \sigma^r \int_0^t \hat{B}(s,t) e^{-\Lambda_{st}^y} ds,$$

and using the results from the Appendix on Hull-White model, the solution to Equations 4.14 and 4.15 are

$$y_t = y_s e^{-\Lambda_{st}^y} - \sigma^{r2} \int_s^t \hat{B}(u,t) e^{-\Lambda_{ut}^y} du + \sigma^r \int_s^t e^{-\Lambda_{ut}^y} dW_u^{rQ_t},$$
$$z_t = z_s e^{-\Lambda_{st}^z} + \nu \int_s^t \zeta_u e^{-\Lambda_{ut}^z} du + \lambda \int_s^t e^{-\Lambda_{ut}^z} dW_u^{zQ_t},$$

where

$$\Lambda_{st}^y = k(t-s), \qquad \Lambda_{st}^z = \nu(t-s).$$

The stock price can then be recovered by:

$$S(t_{i+1}) = S(0) \frac{\exp(x(t_{i+1}) - d(t_{i+1}))}{P(0, t_{i+1})}.$$

For completeness, a second-order accurate method is now described as seen in Glasserman [13] which was implemented to reduce variance with larger steps taken in the PDE solver and simulation. Define a set of d stochastic processes where process x_n is given by

$$dx_n(t) = a_n(x(t))dt + \sum_{k=1}^m b_{nk}(x(t)) \Delta W^k(t),$$

with some correlation structure. Then the discretization for process x_n, with $h = (t_{i+1} - t_i)$, is defined to be:

$$x_n(t_{i+1}) = x_n(t_i) + a_n h + \sum_{k=1}^m b_{nk} \Delta W^k(t_{i+1}) + \frac{1}{2} L^0 a_n h^2 + \frac{1}{2} \sum_{k=1}^m (L^k a_n + L^0 b_{nk}) \Delta W^k(t_{i+1})$$
$$\frac{1}{2} \sum_{k=1}^m \sum_{j=1}^m L^j b_{nk} (\Delta W^j(t_{i+1}) \Delta W^k(t_{i+1}) - V_{jk}),$$

where the differential operators are defined as

$$L^0 = \frac{\partial}{\partial t} + \sum_{n=1}^{d} a_n \frac{\partial}{\partial x_n} + \frac{1}{2} \sum_{n,l=1}^{d} \sum_{k=1}^{m} b_{nk} b_{lk} \frac{\partial^2}{\partial x_n \partial x_l},$$

and

$$L^k = \sum_{n=1}^{d} b_{nk} \frac{\partial}{\partial x_n},$$

for $k = 1, ...m$. Also, $V_{jk} = V_{kj}$ and is defined as a random variable taking values of $-h$ and h with probability of 1/2 each; define $V_{jj} = h$.

Now this second-order scheme is only implemented for the stock return process, since marginal distributions for the interest rate and the random factor z are known exactly. So, consider the system of SDEs in Equations 4.14-4-16 in a slightly different form:

$$dx_t = \left(g_t(y_t) - \frac{1}{2} \sigma_{*t}^{s2} e^{2z_t} + \rho \sigma^r \sigma_{*t}^{s} e^{z_t} \hat{B}(t,T) \right) dt + \sigma_{*t}^{s} e^{z_t} dW_t^{sQ_T}, \quad (4.22)$$

$$dy_t = -\left(k y_t + \sigma^{r2} \hat{B}(t,T) \right) dt + (\sigma_1 dW_t^{rQ_T} + \sigma_2 dW_t^{rQ_T}), \quad (4.23)$$

$$dz_t = \nu(\zeta_t - z_t) dt + \lambda dW_t^{zQ_T}, \quad (4.24)$$

where the Brownian motions are pairwise independent, $\sigma^{r2} = \sigma_1^2 + \sigma_2^2$, and $\sigma_1 = \sigma^r \rho$ (with ρ being the original correlation between stock and interest rate).

In the context of the SR-SLV model for this method we have:

$$a_1 = g_t(y_t) - \frac{1}{2} \sigma_{*t}^{s2} e^{2z_t} + \rho \sigma^r \sigma_{*t}^{s} e^{z_t} \hat{B}(t,T),$$

$$a_2 = -k y_t - \sigma^{r2} \hat{B}(t,T),$$

$$a_3 = \nu(\zeta_t - z_t),$$

$$b_{11} = \sigma_{*t}^{s} e^{z_t}, \quad b_{12} = 0, \qquad b_{13} = 0,$$

$$b_{21} = \sigma_1, \qquad b_{22} = \sigma_2, \qquad b_{23} = 0,$$

$$b_{31} = 0, \qquad b_{23} = 0, \qquad b_{33} = \lambda.$$

Then the second-order discretization for x_t or $x_1 = x$ becomes:

$$x(t_{i+1}) = x(t_i) + a_1 h + b_{11}\Delta W^s(t_{i+1}) + \frac{1}{2}L^0 a_1 h^2 + \frac{1}{2}\sum_{k=1}^{3} L^k a_n + L^0 b_{11}\Delta W^s(t_{i+1})$$
$$\frac{1}{2}\sum_{j=1}^{3} L^j b_{11}(\Delta W^j(t_{i+1})\Delta W^s(t_{i+1}) - V_{j1}),$$

where $W^1 = W^s$, $W^2 = W^r$, and $W^3 = W^z$. After some work (see Appendix) the scheme simplifies to:

$$x(t_{i+1}) = x(t_i) + A\Delta W_1 + B\Delta W_2 + C\Delta W_3 + D$$
$$+\frac{b_{11}}{2}\Big[e^z\frac{\partial\sigma_*^s}{\partial x}(\Delta W_1^2 - h) + b_{33}(\Delta W_3\Delta W_1 + \xi)\Big],$$

where

$$A = \frac{h}{2}\Big[b_{11}\Big(-\frac{e^{2z}}{2}\frac{\partial\sigma_*^{s2}}{\partial x} + \rho\sigma^r\frac{\partial\sigma_*^s}{\partial x}e^z\hat{B}\Big)$$
$$+\Big[a_1 e^z\frac{\partial\sigma_*^s}{\partial x} + a_3\sigma_*^s e^z + \frac{1}{2}\Big(b_{11}^2 e^z\frac{\partial^2\sigma_*^s}{\partial x^2} + \lambda^2\sigma_*^s e^z\Big)\Big]\Big] + b_{11} + \frac{h}{2}b_{21},$$

$$B = b_{22},$$

$$C = \lambda\Big(-b_{11}^2 + \rho\sigma^r\sigma_*^s e^z\hat{B}\Big),$$

$$D = a_1 h + \frac{h^2}{2}\Big[a_1\Big(-\frac{e^{2z}}{2}\frac{\partial\sigma_*^{s2}}{\partial x} + \rho\sigma^r\frac{\partial\sigma_*^s}{\partial x}e^z\hat{B}\Big) + a_2 + a_3\Big(-b_{11}^2 + \rho\sigma^r\sigma_*^s e^z\hat{B}\Big)$$
$$-\rho\sigma^r\sigma_*^s e^{-kh} + \frac{1}{2}\Big(b_{11}^2\Big(-\frac{e^{2z}}{2}\frac{\partial^2\sigma_*^{s2}}{\partial x^2} + \rho\sigma^r\frac{\partial^2\sigma_*^s}{\partial x^2}e^z\hat{B}\Big)$$
$$+\lambda^2\Big(-2b_{11}^2 + \rho\sigma^r\sigma_*^s e^z\hat{B}\Big)\Big],$$

where ξ is a random variable independent of the Brownian motions taking values of $-h$ and h with probability of .5 each. And, to simplify things, we can let

$$\frac{\partial\sigma_*^{s2}}{\partial x} = 2\sigma_*^s\frac{\partial\sigma_*^s}{\partial x},$$

and

$$\frac{\partial^2 \sigma_*^{s2}}{\partial x^2} = 2\left(\frac{\partial \sigma_*^s}{\partial x} + \sigma_*^s \frac{\partial^2 \sigma_*^s}{\partial x^2}\right).$$

The above method was first introduced by Milstein [13] and it is weak order 2 convergent. Conditions for this include uniformly bounded derivatives of coefficients a_i and b_{ij}. Here, local volatility increases very quickly as the stock price falls to zero. Thus, similarly as in the SR-LV case, it is necessary to bound not only local volatility, but its first two derivatives for convergence and realism.

4.5 Pricing results

This section discusses the numerical results for the pricing of variance/volatility derivatives. As a general guideline, several measures are taken to ensure stability of the PDE solver and low variance of the simulation results.

As discussed in chapter 3, in both SR-LV and SR-SLV models, θ in the Graig-Sneyd scheme governs stability. To insure stability, θ in both models is set to 1. Furthermore, dt/Δ^2 is kept to less than 1/2, where Δ is the space increment size for the uniform mesh. This is an extra precaution, since the θ restiction is only discussed in the constant coefficient case and in our models, coefficients are time and space dependent. Furthermore, the coefficients are unknown and only revealed one step at a time.

Time step of .025 years is used for the simulations with maturities up to 3 years. Variance is reduced by running 100,000 simulations per price. Simulation is relatively inexpensive in computation time (e.g., 1 year variance swap price in about 7 seconds on a home PC) and this number of simulations ensures 4 significant digits of accuracy. Time step can be increased to .05 years for longer maturities, but appropriate adjustment need to be made to the mesh and number of simulations to ensure stability of the PDEs and low variance.

In the simulations, a cubic spline is utilized on the local volatility surface to pick up appropriate volatilities for the simulated stock prices. The coefficients in the simulation must be uniformally bounded and this is a good way to provide a bound at each time. At each time step, if the stock price ventures outside of ± 2.5 standard deviations (approximately) from the mean of the stock price level at that time, then the volatility is taken to be that on this boundary. This prevents extrapolation beyond the local volatility surface (which is about twice as steep as the implied volatility surface) and ensures realistic values for volatility. The restriction on variance is extended in the SR-SLV model where the local volatility bounds are applied to the calibrated leverage function.

The options surface for S&P 500 is used, giving option prices for any maturity and strike by the VGSSD calibrated/stylized model with parameters: $\theta = -0.16782$, $\nu = 0.65289$, $\sigma = 0.19732$, and $\gamma = 0.6$.

All variance swap prices are quoted in volatility.

4.5.1 Verification of modeling via options

Before proceeding to pricing variance swaps, it would be prudent to test the modeling environment. The SR-LV model extracts the local volatility surface in the stochastic interest rate environment. Then the SR-SLV model calibrates the leverage function to this surface. However, can the SR-LV model replicate the option prices given by the VGSSD model? To test the validity of the local volatility surface, 5 option prices were generated at 3 different maturities using the SR-LV model and the results are compared to the prices given by the VGSSD model.

The following table presents prices generated for June 19, 2002 by 100,000 simulations using the calibrated parameters of the VGSSD and Hull-White models. Since the forward density is known at all times, the options can be priced by integrating over the density at the appropriate time. However, pricing via simulation adds another check of consistency. Assumed correlation between S&P 500 and the short rate for this table is zero. The prices are given for the SR-LV model with the VGSSD price in parenthesis. The spot price is taken to be \$1000 and the strikes are \$800, \$900, \$1000, \$1100, and \$1200. Put prices are given for the first three strikes and call prices are given for the latter two strikes, so that we observe the out-of-the-money pricing performance.

T	K=\$800	K=\$900	K=\$1000	K=\$1100	K=\$1200
1yr	20.14 (21.691)	41.47 (42.87)	77.23 (77.56)	43.95 (43.83)	19.05 (18.31)
2yr	39.17 (41.53)	65.33 (67.34)	102.10 (102.90)	93.13 (92.89)	59.32 (57.76)
3yr	51.51 (53.96)	78.61 (80.68)	114.13 (114.97)	137.11 (138.08)	100.78 (99.77)

Table 4.1: Out-of-the-money options with spot price of \$1000 and strikes \$800, \$900, \$1000, \$1100, and \$1200 (put taken at strike \$1000) are priced using VGSSD and SR-LV model on June 19, 2002.

The average absolute error for the 15 options is 2.5%. Deep out-of-the-money puts produce most inaccuracy. This is a consequence of bounding the local volatility during simulation. When the bounds are increased, the put prices generated by the SR-LV model approach to the corresponding VGSSD prices. Thus, the local volatility surface in SR-LV setting is consistent with the VGSSD option surface.

4.5.2 Variance swaps

SR-LV and SR-SLV models will give the same price for the variance swap. Thus, only the SR-LV model is simulated where the realized variance is computed along each path and averaged over the 100,000 simulations. Two effects are discussed in this section and compared to the pricing of the variance swaps via options replication of the log contract. First, the correlation structure is examined for 1, 2, and 3 year swaps. Then, the term structure of variance swaps is discussed for several levels of volatility of the short rate. In general, intuition developed in section 4.1 holds with the observed results.

Table 4.2 represents variance swap prices for various levels of correlation (between $\pm .5$). There is little evidence in the literature that the stock price and short rates are strongly correlated. Many studies find weak correlation at best. Average computation time on a home PC (includes PDE solver and simulations) for 1-year price is 26 seconds, 2-year - 53 seconds, and 3-year - 79 seconds. The default calibrated parameters for the short rate model are $k = 0.025091$ and $\sigma^r = 0.011591$.

As expected, when correlation increases (in the positive direction), variance swap price falls. This relation is almost linear in correlation. For a negative enough

ρ	-.5	-.4	-.3	-.2	-.1	0	.1	.2	.3	.4	.5
1yr K_{vs}	22.46	22.41	22.36	22.27	22.21	22.13	22.07	22.00	21.93	21.86	21.79
2yr K_{vs}	24.02	23.87	23.73	23.59	23.45	23.30	23.16	23.04	22.90	22.74	22.62
3yr K_{vs}	25.18	24.99	24.75	24.56	24.39	24.15	23.96	23.77	23.56	23.37	23.19

Table 4.2: Variance swap prices for 1, 2, and 3 years for different levels of correlation. The variance swap prices via replication of the log contract is 22.9 for 1 yrear, 24.06 for 2 years, and 24.69 for 3 years.

level of correlation, variance swap price may even exceed the price given by the log-contract hedge. Also, one would expect stochastic interest rates to play a more important role as the maturity of the variance swap contract increases. That is, the difference from the log-contract hedge price should increase in time. This does not seem to be the case here. As maturity increases, the SR-LV model price approaches the log-contract hedge price.

Observe that the calibrated volatility of the short rate is fairly low - 0.011591. Thus, the short rate does not vary much and the results therefore resemble the pattern set by the log-contract hedge price. Table 4.3 shows the term structure of variance swaps for varying levels for the short rate volatility when the correlation between the stock and short rate is zero. When the short rate volatility is .04, we begin to see the expected pattern in the difference from the log-contract hedge price: the difference between SR-LV model and log-contract hedge price increases with maturity. More interestingly, the increasing pattern of variance swap prices via option replication of the log contract is broken in the SR-LV setting when the short rate volatility is .07. Thus, stochastic interest rate play a more important role in the pricing of variance swaps as the short rate volatility increases.

Recent economic events suggest that the pattern of little correlation between

$Time$	1	1.5	2	2.5	3
K_{lc}	22.90	23.59	24.06	24.41	24.69
K_{vs} $\sigma^r = .0116$	22.13	22.78	23.30	23.71	24.15
K_{vs} $\sigma^r = .04$	22.06	22.53	22.87	22.92	23.14
K_{vs} $\sigma^r = .07$	21.81	21.99	21.82	21.28	20.85

Table 4.3: Variance swap prices for 1-3 years for varying short rate volatility levels - .0116 (calibrated), .04,.07, and $\rho = 0$.

equities and interest rates may be changing. That is, the Federal Reserve seems to be lowering the funds rate whenever the equity markets fall creating correlation between equities and the interest rates. Thus, there is a need to examine pricing under higher correlations. For completeness, Table 4.4 shows 1-year variance swap prices for complete range of correlation. Correlation of ± 1 is omitted because there is a discontinuity in the initial distribution at those values.

ρ	-.9	-.8	-.7	-.6	-.5	-.4	-.3	-.2	-.1	0
1yr K_{vs}	22.73	22.67	22.61	22.55	22.46	22.41	22.36	22.27	22.21	22.13
ρ	.1	.2	.3	.4	.5	.6	.7	.8	.9	
1yr K_{vs}	22.07	22.00	21.93	21.86	21.79	21.74	21.68	21.57	21.54	

Table 4.4: Variance swap prices for 1 year for expanded levels of correlation. The variance swap price via replication of the log contract is 22.9 for 1 yrear.

4.5.3 Variance options

Variance options can be priced using the SR-SLV model which introduces the volatility of volatility (vol of vol) component - λ. The SR-LV model only produces one price for a given option because it lacks vol of vol. Madan *et al.* [19] comments that vol of vol is poorly calibrated from the options surface and this dissertation will focus on analyzing option prices for various levels of vol of vol and on mean reversion rate of the stochastic component of volatility.

In this sections, staddles will be priced. A straddle is a position in a put and a call together bought at the same strike. Trading a straddle is equivalent to trading volatility. If an investor buys a straddle, he/she is hoping for volatility as payoffs occur when the underlying asset moves away from strike in either direction. Selling a straddle is the opposite case - betting that the underlying asset will not move much to maturity. In the context of this dissertation, the value of straddles on realized variance depends on the movement of realized variances for some maturity.

Tables 4.4 to 4.6 represent straddle prices for the S&P 500 on 1-year realized variance using calibrated paremeters for the short rate model: $k = 0.025091$ and $\sigma^r = 0.011591$ with a strike variance of .04. Average computational time for one straddle price was 54 seconds. This time includes one PDE solver recovering the local volatility surface, one PDE solver calibrating the leverage function, and simulations. The three tables differ by the correlation level between the short rate and stock. The LV column corresponds to the straddle price in the SR-LV framework, or when the $\lambda \to 0$ in the SR-SLV model.

Note that the straddle prices increase with vol of vol (λ) and decrease with the mean reversion rate (ν). It seems that the level of correlation does not play a big role in straddle prices on 1-year realized variance. However, this may be a consequence of the relatively low level of the calibrated short rate volatility and a maturity of just a year.

To examine how straddle prices depend on volatility of the short rate and cor-

$\rho = 0$		λ		
ν	LV	25%	50%	75%
.5	1.42	1.69	2.28	2.97
1	1.42	1.62	2.09	2.67
2	1.42	1.54	1.84	2.21
4	1.42	1.47	1.61	1.79

Table 4.5: 1-year variance option prices for different levels of mean reversion rate (ν) and vol of vol (λ) for zero correlation between stock and interest rate.

$\rho = .5$		λ		
ν	LV	25%	50%	75%
.5	1.36	1.62	2.21	2.90
1	1.36	1.56	2.01	2.55
2	1.36	1.47	1.76	2.14
4	1.36	1.40	1.53	1.72

Table 4.6: 1-year variance option prices for different levels of mean reversion rate (ν) and vol of vol (λ) for .5 correlation between stock and interest rate.

relation when maturity is increased to three years, straddle prices were generated and results are summarized in Table 4.7. For a given volatility of the short rate, straddle price decreases with an increase in correlation (in the positive direction). When the stock price is positively correlated with the short rate, then the straddle price falls as volatility of the short rate increases. Under negative correlation, this phenomenon is reversed. This is an indication that with an increase in volatility of the short rate, the slope between straddle price and correlation becomes more negative. Thus, the magnitude of the impact of correlation on straddle prices depends on time to maturity and volatility of the short rate.

Chapter 3 discussed a stylistic feature of models with stochastic volatility. For completeness, plots of the 1-year marginal distribution in stock (x) and short rate (y) for large z and different inputs of ν and λ. When ν is large and λ is small (Figure

$\rho = -.5$		λ		
ν	LV	25%	50%	75%
.5	1.49	1.77	2.38	3.07
1	1.49	1.70	2.17	2.70
2	1.49	1.62	1.92	2.26
4	1.49	1.54	1.69	1.87

Table 4.7: 1-year variance option prices for different levels of mean reversion rate (ν) and vol of vol (λ) for $-.5$ correlation between stock and interest rate.

	ρ				
σ^r	-.5	-.25	0	.25	.5
.0116	3.47	3.31	3.16	2.98	2.80
.04	4.16	3.41	2.85	2.45	2.21

Table 4.8: 3-year variance straddle prices across correlations and short rate vol for $\lambda = .5$ and $\nu = .5$

4.3), the SR-SLV model behaves more like the SR-LV model and the distribution is unimodal. As λ increases, the distribution becomes more bimodal. Since the factor z represents a random shock to volatility, the market will likely find another support level signified by the second mode. Mathematically, large z results in large variance which involves a squared of the price change. The two modes reflect that the large square of the price change could have come from a big upward or downward jump in the price. The bimodal feature is most evident when λ is high and ν is low for large shock in z.

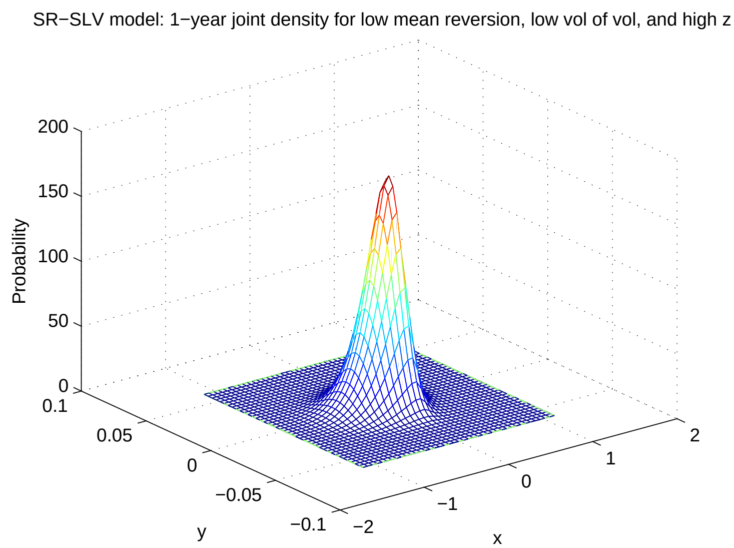

Figure 4.1: SR-SLV model: View of 1-year marginal density of x and y when z is large and $\nu = .5$ and $\lambda = .05$.

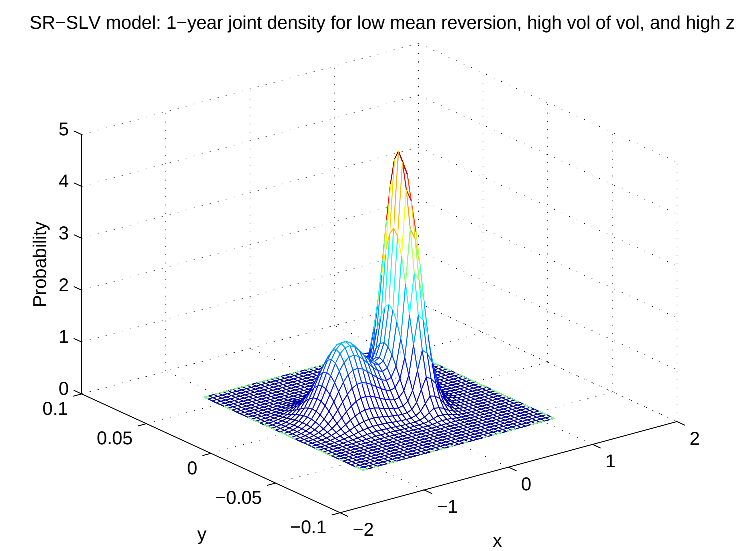

Figure 4.2: SR-SLV model: View of 1-year marginal density of x and y when z is large and $\nu = .5$ and $\lambda = .75$. Zero correlation between stock and interest rate is assumed.

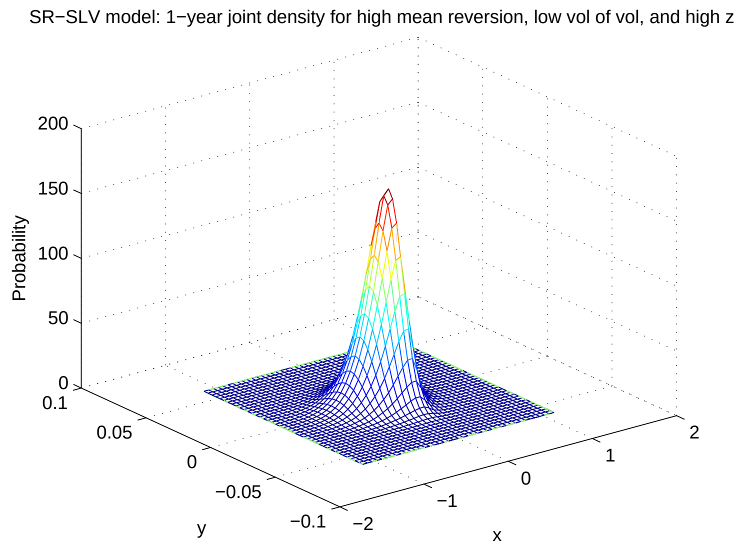

Figure 4.3: SR-SLV model: View of 1-year marginal density of x and y when z is large and $\nu = 4$ and $\lambda = .05$. Zero correlation between stock and interest rate is assumed.

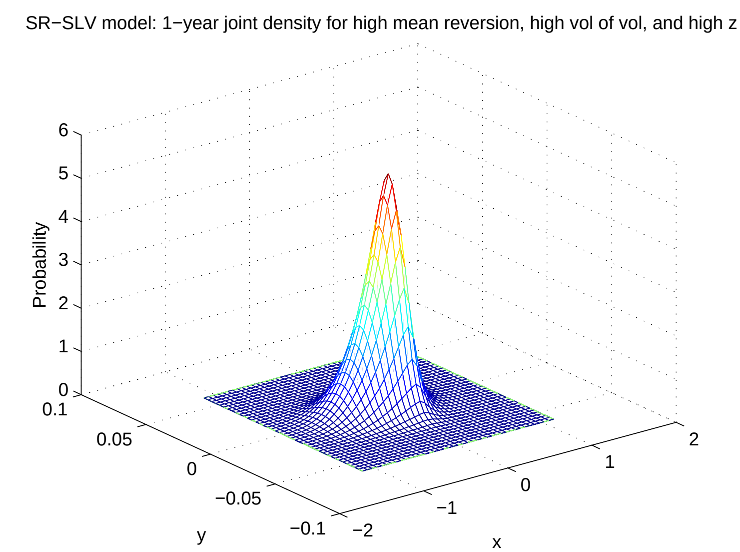

Figure 4.4: SR-SLV model: View of 1-year marginal density of x and y when z is large and $\nu = 4$ and $\lambda = .75$. Zero correlation between stock and interest rate is assumed.

4.5.4 Volatility swaps/options

Simulation also works well and is a very natural methodology for pricing volatility swaps. The volatility swap has a similar payout as the variance swap contract, only it is on volatility - square root of variance. We can just talk about the pricing of this contract in the SR-LV model. That value of the volatility swap at inception is zero, setting the strike in the contract to:

$$K_{vol} = E^{Q_T}\left[\int_0^T \sigma^s dt\right],$$

where $\sigma^s = \sigma^s(S_t, t)$. For reference, the fair strike for the variance swap is

$$K_{var} = E^{Q_T}\left[\int_0^T \sigma^{s2} dt\right].$$

Table 4.8 shows volatility swap prices with 3-year maturity across different correlation for S&P 500. The calibrated short rate parametes are used in the simulations. Same as in the previous section, 100,000 simulations are run per price to ensure low variance. Variance swap prices are quoted in volatility.

Observe that the volatility swap price is lower than the variance swap (in vol). For variance swaps, variances across paths are summed, averaged, and the result is square rooted. For volatility swaps, volatility across paths are summed and averaged. The validity of the lower volatility swap price can easily be checked using the identity - $(a^2 + b^2 - 2ab) \geq 0$.

Table 4.9 gives 1-year volatility straddle prices for zero correlation between stock price and short rate. Same conclusions hold across different levels of speed of mean reversion and vol of vol as did for variance swaps. An interesting observation

ρ	-.5	-.4	-.3	-.2	-.1	0	.1	.2	.3	.4	.5
K_{vol}	24.29	24.06	23.90	23.74	23.52	23.34	23.16	22.98	22.79	22.62	22.39
K_{var}	25.18	24.99	24.75	24.56	24.39	24.15	23.96	23.77	23.56	23.37	23.19

Table 4.9: 3-year volatility swap prices and variance swap prices for different levels of correlation.

is that volatility straddles are a bit more than twice the price of the variance swaps for the same strike (in vol). This is a consequence of the payout being about twice as large for the volatility straddle. The following proposition outlines the details.

$\rho = 0$		λ		
ν	LV	25%	50%	75%
.5	2.97	3.56	4.83	6.20
1	2.97	3.41	4.41	5.61
2	2.97	3.26	3.86	4.69
4	2.97	3.10	3.36	3.78

Table 4.10: 1-year volatility option prices for different levels of mean reversion rate (ν) and vol of vol (λ) for zero correlation between stock and interest rate.

Proposition The price of a volatility straddle is about twice the price of a variance straddle with the same strike (in terms of vol).

Proof: A straddle has the same payoff as a put and a call options together. That is, suppose some asset has a price X_t and strike level K_X, then the payoff at time T of a straddle is given by $S(X) = (X_T - K_X)^+ + (K_X - X_T)^+$. Now, consider expansions of variance and volatility straddles around the strike to the first order:

$$S(\sigma) = S(K_\sigma) + \frac{\partial S}{\partial \sigma}(\sigma - K_\sigma)$$

$$S(\sigma^2) = S(K_{\sigma^2}) + \frac{\partial S}{\partial \sigma^2}(\sigma^2 - K_{\sigma^2})$$

But,

$$\frac{\partial S}{\partial \sigma} = \frac{\partial S}{\partial \sigma^2}\frac{\partial \sigma^2}{\partial \sigma} = \frac{\partial S}{\partial \sigma^2} 2\sigma.$$

Thus, the two payoffs become:

$$S(\sigma) = S(K_\sigma) + \frac{\partial S}{\partial \sigma^2} 2\sigma(\sigma - K_\sigma) \tag{4.25}$$

$$S(\sigma^2) = S(K_{\sigma^2}) + \frac{\partial S}{\partial \sigma^2}(\sigma^2 - K_{\sigma^2}). \tag{4.26}$$

$S(K_\sigma)$ and $S(K_{\sigma^2})$ are payoffs at strike and for the straddles this is zero. In the context of earlier discussion, $K_\sigma = K_{vol}$ and $K_{\sigma^2} = K_{var}$. Dividing 4.25 by 2 and subtractinig 4.26 from 4.25 results in:

$$\frac{S(\sigma)}{2} - S(\sigma^2) = \frac{\partial S}{\partial \sigma^2}(K_{\sigma^2} - \sigma K_\sigma).$$

Let $K_{\sigma^2} = K_\sigma^2$, so that the straddles are considered for the same volatility strike. Then we have

$$S(\sigma) = 2S(\sigma^2) + 2\frac{\partial S}{\partial \sigma^2}(K_\sigma^2 - \sigma K_\sigma).$$

Therefore, for the same strike level (in terms of vol) the payoff of the volatility straddle is twice the payoff of the variance straddle plus some convexity adjustment. Thus, the price of the volatility straddle should be at least twice the price of a variance straddle with the same strike (in vol).

4.6 Replication of variance swaps under stochastic interest rates

Realized variance from time 0 to T is given by

$$RV(0,T) = \frac{2}{T}\left[\int_0^T \frac{dS_u}{S_u} - \ln\frac{S_T}{S_0}\right]. \tag{4.27}$$

The replication of the log contract by a portfolio of options is the subject of Chapter 2 and this section focuses on the dynamic component of realized variance, $\Pi_T = \int_0^T \frac{dS_u}{S_u}$. An approximate replicating strategy is developed for Π_T independent of model assumptions and the error is explored in the SR-LV model setting.

Suppose the interest rate is deterministic. The instantaneous return, $\frac{dS_t}{S_t}$, is realized at time t when holding $\frac{1}{S_t}$ shares, but it must be paid at maturity, T. This return can be deposited in a bank where it will earn the risk-free rate of interest from time t to T. Thus, if $\frac{dS_t}{S_t}$ is desired at maturity, then the appropriate number of shares that need to be held at t is $\exp(-\int_t^T r_s ds)/S_t$. Therefore, there exists a replicating strategy for Π_T involving positions in the stock and money-market account when the interest rate is deterministic.

When the interest rate is stochastic, the amount of cash deposited in the bank is $\exp(-\int_t^T r_s ds)\frac{dS_t}{S_t}$ is a function of two random variables, S_t and r_t. Furthermore, the amount of interest earned from t to T is also a random quantity. With the money-market account, the positions in the above replication are no longer predictable. Instead, invest the return at time t in a zero-coupon bond expiring at time T, $P(t,T)$. For ease of notation, let $\tilde{S}_t = S_t/P(t,T)$ and $\Pi_T^* = \int_0^T \frac{d\tilde{S}}{\tilde{S}}$. Consider the following strategy that replicates Π_T^*:

$$
\begin{aligned}
V_0^* &= 0,\\
\alpha_t^* &= \frac{1}{\tilde{S}_t},\\
\beta_t^* &= \int_0^t \frac{1}{\tilde{S}_u} d\tilde{S}_u - 1,\\
V_t^* &= V_0^* + \alpha_t^* S_t + \beta_t^* P(t,T),
\end{aligned}
$$

where α_t^* and β_t^* are predictable positions in the stock and bond respectively. At maturity, $V_T^* = \Pi_T^*$:

$$
V_T = V_0 + \alpha_T^* S_T + \beta_T^* P(T,T) = \frac{1}{\tilde{S}_T} S_T + \left(\int_0^T \frac{1}{\tilde{S}_u} d\tilde{S}_u - 1 \right) P(T,T) = \int_0^T \frac{1}{\tilde{S}_u} d\tilde{S}_u.
$$

However, for $t < T$ we have $V_t^* < \Pi_t^*$:

$$
V_t = V_0 + \alpha_t^* S_t + \beta_t^* P(t,T) = \frac{1}{\tilde{S}_t} S_t + \left(\int_0^t \frac{1}{\tilde{S}_u} d\tilde{S}_u - 1 \right) P(t,T) = P(t,T) \int_0^t \frac{1}{\tilde{S}_u} d\tilde{S}_u < \Pi_T^*.
$$

This replicating strategy is also self-financing. Observe the P&L (let $t_1 < t_2$, $t_2 - t_1$ small) for the stock and bond position respectively:

$$
\begin{aligned}
\left(\frac{1}{\tilde{S}_{t_2}} - \frac{1}{\tilde{S}_{t_1}} \right) S_{t_2} &= P(t_2,T) - P(t_1,T) \frac{S_{t_2}}{S_{t_1}},\\
\left(\frac{\tilde{S}_{t_2}}{\tilde{S}_{t_1}} - 1 \right) P(t_2,T) &= P(t_1,T) \frac{S_{t_2}}{S_{t_1}} - P(t_2,T).
\end{aligned}
$$

Dynamically trading the above portfolio only guarantees Π_T^* at time T, but $\Pi_T^* \neq \Pi_T$. To examine the error, assume SR-LV model setting under the risk-neutral measure. Using Itô's quotient rule we get

$$
\Pi_T^* = \int_0^T \frac{d\tilde{S}}{\tilde{S}} = \int_0^T \left[(-q + \sigma^{r2} \hat{B}^2(t,T) - \rho \sigma^r \sigma^s \hat{B}(t,T)) dt + \sigma^r \hat{B}(t,T) dW^r + \sigma^s dW^s \right], \tag{4.28}
$$

where

$$\frac{dP(t,T)}{P(t,T)} = r_t dt - \sigma^r \hat{B}(t,T) dW^r,$$

and $\hat{B}(t,T) = (1-e^{-k(T-t)})/k$ as before. Substituting for $\sigma^r \hat{B}(t,T)dW^r$ in Equation 4.28 using the bond SDE results in

$$\Pi_T^* = \int_0^T \left[(r_t - q + \sigma^{r2} \hat{B}^2(t,T) - \rho\sigma^r\sigma^s \hat{B}(t,T))dt + \sigma^s dW^s - \frac{dP(t,T)}{P(t,T)} \right]. \quad (4.29)$$

Now, the aim is to replicate Π_T which is given by:

$$\Pi_T = \int_0^T (r_t - q)dt + \sigma^s dW^s. \quad (4.30)$$

Subtracting Equation 4.30 from 4.29 gives the error, E:

$$E = \Pi_T^* - \Pi_T = \int_0^T (\sigma^{r2} \hat{B}^2(t,T) - \rho\sigma^r\sigma^s \hat{B}(t,T))dt - \int_0^T \frac{dP(t,T)}{P(t,T)}.$$

The first observation is that the error contains a stochastic component (last term), which means that this is not a hedge, but an approximate replication for Π_T. Secondly, when σ^r is relatively small, the error is expected to be small as well. Finally, the replicating portfolio for Π_T^* is composed of stock and bond positions, so if the stock and the interest rate are correlated, the portfolio is expected to reflect this relationship. However, from the error it can be seen that some correlation is still missing. Therefore, for this replication to perform well for Π_T, the volatility of the short rate cannot be too large.

Equation 4.27 with the approximate replication was implemented in the SR-LV model context for 1-year variance swaps using simulation under the T-forward measure. In the Hull-White model, $P(t,T)$ is known in a closed form (Overhaus *et* *al.* [23]):

$$P(t,T) = \frac{P(0,T)}{P(0,t)} \exp\left(- (r_t - f(0,t))\hat{B}(t,T) - \frac{1}{2}\hat{B}(t,T)^2 \sigma^{r2} \int_0^t e^{-2k(t-s)} ds \right),$$

where $r_t - f(0,t) = \sigma^r \int_0^t e^{-k(t-s)} dW_s^{Q_t}$.

Pricing variance swaps via replication is extremely accurate when compared to the SR-LV model prices for varying levels of model parameters. However, deviations are observed on individual simulated paths. Consider the calibrated Hull-White parameters for the SR-LV model on June 19, 2002: $k = 0.025$ and $\sigma^r = 0.012$. Errors are reported for 100,000 simulations of 1-year SR-LV model, where the absolute difference is taken between the variance given by the replication and variance computed by the sum of squared log differences on each simulated path.

For $\rho = 0$, the average absolute error of the replicating strategy from actual realized variance was .00049. For $\rho = .8$, the error was .00045. This means, if the realized variance on some path was $(20\%)^2$ then the replication may give $(20.11\%)^2$. Thus, changing correlation level does not seem to impact the accuracy of the replication.

The rate volatility, σ^r, was increased to .04 and k stayed at .025. The absolute error increased to .00113 and remained near this level as correlation was increased from 0 to .8. As expected from the error analysis, the error increases with a rise in short rate volatility. Better replicating strategies is a direction for future research.

Chapter 5

Conclusions

5.1 General results

Some conventional methods exist for pricing variance derivatives (e.g., variance swaps) using portfolios of options traded in the market. This approach overprices variance derivatives since an option prices in the variance of the interest rates as well. The work in this dissertation focuses on pricing variance derivatives by correctly accounting for the stochasticity of the interest rates. This is accomplished by two models: SR-LV and SR-SLV. Both models utilize György's [14] result on deriving a Markov process from some n-dimensional Itô process.

In the SR-LV model, stock return follows a diffusion, local volatility is a deterministic function of stock price and time. The short rate is stochastic and driven by the Hull-White process. A closed form expression for local volatility in this framework is available and can be evaluated under the T-forward measure using the joint density of stock and short rate. The local volatility surface is derived while simultaneously solving the appropriate PDE for the joint density.

In the SR-SLV model, stock return also follows a diffusion, the short rate is stochastic, volatility is governed by a leverage function and another random factor uncorrelated with the return of the stock and the short rate. At times, volatility in the market jumps without the stock price moving. This feature is captured in the

SR-SLV model, but not in the SR-LV model. The leverage function is calibrated to the local volatility surface while simultaneously solving the PDE for the joint density of the stock, short rate, and the random component of volatility. The marginal distribution of the stock price was observed to be bimodal for large levels of the random component of volatility. This is a stylistic feature that should be present in models involving stochastic volatility. Intuitively, the effect shows that with a large random jump in volatility, the market finds another support level at the second mode.

Simulation is a natural approach to value variance derivatives utilizing the cubic spline on the local volatility surface. In this dissertation variance swaps are priced for different maturities and across different levels of correlation between the stock and the short rate. The SR-LV model produces prices which are lower than the conventional method of pricing by replicating the log contract with options. Furthermore, it is shown that the difference from the conventional method increases as the volatility of the short rate increases. Straddles on realized variance are priced using the SR-SLV model for different maturities and levels of the speed of mean reversion and volatility of the random component of stock volatility. As expected, staddle prices increase with volatility of volatility and decrease with the speed of mean reversion. The slope between the straddle prices and correlation also becomes more negative with an increase in volatility of the short rate.

The work in this dissertation is extended to volatility derivatives. For exposition, 3-year volatility swap prices were generated across different levels of correlation. Volatility swap prices are lower than the corresponding prices of the variance swap prices (in volatility). Straddle prices on realized volatility are calculated and it was observed that those were about twice the price of the straddle prices on realized variance with the same strike (in volatility). It is shown that this is the consequence of the payoff of the straddle on volatility being twice as big as the payoff for the corresponding variance straddle.

5.2 Future work

This dissertation answered questions regarding correct pricing of variance derivatives in the presence of stochastic interest rates. However, in addition to knowing the hedge, a trader may want to know the risk (P&L (Profit and Loss) sensitivity to factors such as stock price movements, short rate volatility changes, correlation, etc) associated with the position.

Focusing on variance swaps in the risk-neutral setting, the fair variance strike under deterministic rates is given by

$$K = \frac{2}{T} E\left[\int_0^T \frac{dS}{S} - \ln\frac{S_T}{S_0}\right] = \frac{2}{T}\left[\int_0^T r_t dt - qT - E\left(\ln\frac{S_T}{S_0}\right)\right]. \tag{5.1}$$

Equation 5.1 reveals that the expected realized variance can be decomposed into a long dynamic position in the stock and a short static position in the log contract replicated by the portfolio of out-of-the-money options. The cost of dynamic rebalancing is therefore

$$P_D = \int_0^T r_t dt - qT.$$

Under stochastic interest rates and the risk-neutral setting, the fair strike is given by

$$K = \frac{E\Big[e^{-\int_0^T r_t dt} RV(0,T)\Big]}{E\Big[e^{-\int_0^T r_t dt}\Big]}, \tag{5.2}$$

where

$$RV(0,T) = \frac{2}{T}\bigg[\int_0^T \frac{dS}{S} - \ln\frac{S_T}{S_0}\bigg].$$

But, Equation 5.2 is not pleasant for evaluating risk. If an assumption is made that the correlation between the stock price and short rate is zero, then, while not independent, we have

$$K \approx \frac{2}{T}E\bigg[\int_0^T \frac{dS}{S} - \ln\frac{S_T}{S_0}\bigg]. \tag{5.3}$$

In SR-LV context of this thesis we have:

$$\frac{dS_t}{S_t} = (r_t - q)dt + \sigma^s(S,t)dW_t^s,$$

$$dr_t = (\theta_t - kr_t)dt + \sigma^r dW_t^r,$$

$$dW_t^s dW_t^r = \rho dt.$$

Letting $x_t = \ln(S_t/F_t)$ and writing r_t as an O-U process gives:

$$dx_t = \Big(g_t(y_t) - \frac{1}{2}\sigma^{s2}\Big)dt + \sigma^s dW_t^s,$$

$$dy_t = -ky_t dt + \sigma^r dW_t^r,$$

$$dW_t^s dW_t^r = \rho dt.$$

where $g_t(y_t) = r_t - f(0,t) = x_t + \int_0^t \sigma^{r2}\frac{(1-e^{-k(t-s)})}{k}e^{-k(t-s)}ds$ and it was shown how to recover $\sigma^s(S,t)$.

Now everything is finite in Equation 5.3 for interchanging integration and expectation and $\rho = 0$ gives

$$K \approx \frac{2}{T}E\bigg[\int_0^T \frac{dS}{S} - \ln\frac{S_T}{S_0}\bigg] = \frac{2}{T}\bigg[\int_0^T E[r_t - q]dt - E\bigg(\ln\frac{S_T}{S_0}\bigg)\bigg] =$$

$$\frac{2}{T}\left[\int_0^T \left(E[x_t] + \int_0^t \sigma^{r2}\frac{(1-e^{-k(t-s)})}{k}e^{-k(t-s)}ds + f(0,t)\right)dt - qT - E\left(\ln\frac{S_T}{S_0}\right)\right].$$

Remembering that $f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T}$ and $E[x_t] = 0$ (using solution to the x_t process and $x_0 = 0$) we get

$$K \approx \frac{2}{T}\left[-\ln P(0,T) - qT + \frac{\sigma^{r2}}{k}\int_0^T\int_0^t (1-e^{-k(t-s)})e^{-k(t-s)}ds\ dt - E\left(\ln\frac{S_T}{S_0}\right)\right]$$

This simplifies to

$$E[RV(0,T)] = \frac{2}{T}\left[-\ln P(0,T) - qT + \frac{\sigma^{r2}}{4k^3}\left[2kT - e^{-2kT} + 4e^{-kT} - 3\right] - E\left(\ln\frac{S_T}{S_0}\right)\right] \tag{5.4}$$

It can be seen that the cost of rebalancing the dynamic position in stock is now:

$$P_S = E\left[\int_0^T \frac{dS}{S}\right] = -\ln P(0,T) - qT + \frac{\sigma^{r2}}{4k^3}\left[2kT - e^{-2kT} + 4e^{-kT} - 3\right]. \tag{5.5}$$

Now consider partials (risk) of the Equation 5.5 with respect to σ^{r2} and k to determine the impact of stochastic interest rates on the rebalancing price:

$$\frac{\partial P_S}{\partial \sigma^{r2}} = \frac{1}{4k^3}\left[2kT - e^{-2kT} + 4e^{-kT} - 3\right] \geq 0, \tag{5.6}$$

$$\frac{\partial P_S}{\partial k} = -\frac{\sigma^{r2}}{4k^4}\left[4kTe^{-kT} - 2kTe^{-2kT} + 2kT - 9 + 12e^{-kT} - 3e^{-2kT}\right] \leq 0. \tag{5.7}$$

The inequalities are true since $kT \geq 0$ ensures the quantities in the brackets of Equations 5.6 and 5.7 to be greater than zero. So there are now several observations regarding risk of volatility of the short rate and speed of mean reversion moving:

1) As $\sigma^r \to 0$ (and/or $k \to \infty$), the price of rebalancing converges to the price under deterministic rates since under the deterministic rates, the short rate is the forward rate.

2) As the variance of the short rate increases, the price of rebalancing increases (risk given by Equation 5.6).

3) As the speed of mean reversion of the short rate increases, the price of rebalancing decreases (risk given by Equation 5.7).

4) The price of rebalancing increases in time for the short rate component as well.

5) We started with $K \approx E[RV(0,T)]$ under $\rho = 0$ assumption. When $\rho \neq 0$, the picture is more complicated, but we can take a look at the price of variance swaps under the T-forward measure.

In the T-forward setting, the SR-LV model is given by:

$$\frac{dS_t}{S_t} = \Big(r_t - q + \rho\sigma^r\sigma^s(S,t)\frac{(1-e^{-k(T-t)})}{k}\Big)dt + \sigma^s(S,t)dW_t^{sQ_T},$$
$$dr_t = \Big(\theta_t - kr_t - \sigma^{r2}\frac{(1-e^{-k(T-t)})}{k}\Big)dt + \sigma^r dW_t^r,$$
$$dW_t^{sQ_T}dW_t^{rQ_T} = \rho dt.$$

Letting $x_t = \ln(S_t/F_t)$ and writing r_t as an O-U process gives:

$$dx_t = \Big(g_t(y_t) - \frac{1}{2}\sigma^{s2} + \rho\sigma^r\sigma^s(S,t)\frac{(1-e^{-k(T-t)})}{k}\Big)dt + \sigma^s dW_t^{sQ_T},$$
$$dy_t = -\Big(ky_t + \sigma^{r2}\frac{(1-e^{-k(T-t)})}{k}\Big)dt + \sigma^r dW_t^{rQ_T},$$
$$dW_t^{sQ_T}dW_t^{rQ_T} = \rho dt.$$

where again $g_t(y_t) = r_t - f(0,t) = x_t + \int_0^t \sigma^{r2}\frac{(1-e^{-k(t-s)})}{k}e^{-k(t-s)}ds$ and

$$\sigma^{s2}(K,T) = \frac{\frac{\partial C}{\partial T} + q(C - KC_K) + KP(0,T)E^{Q_T}\big[r_T\mathbf{1}_{S_T>K}\big]}{\frac{K^2}{2}\frac{\partial^2 C}{\partial K^2}}.$$

Changing to T-forward measure the price of the variance swap

$$K = \frac{2}{T}E^{Q_T}\bigg[\int_0^T \frac{dS}{S} - \ln\frac{S_T}{S_0}\bigg] = \frac{2}{T}\bigg[E^{Q_T}\bigg(\int_0^T r_t dt\bigg) - qT - E^{Q_T}\bigg(\ln\frac{S_T}{S_0}\bigg)\bigg]. \quad (5.8)$$

Under the T-forward measure $E^{Q_T}[r_t] = f(0,T)$ and again using $f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T}$ we get

$$K_{Q_T} = \frac{2}{T}\left[-\ln P(0,T) - qT + \rho\sigma^r \int_0^T \sigma^s(S,t)\frac{(1-e^{-k(T-t)})}{k}dt - E^{Q_T}\left(\ln\frac{S_T}{S_0}\right)\right]. \tag{5.9}$$

Using Taylor series to first order $\frac{(1-e^{-k(T-t)})}{k} \approx T-t$ and

$$K_{Q_T} = \frac{2}{T}\left[-\ln P(0,T) - qT + \rho\sigma^r \int_0^T \sigma^s(S,t)(T-t)dt - E^{Q_T}\left(\ln\frac{S_T}{S_0}\right)\right]. \tag{5.10}$$

This expression is awkward since the realized variance is written in terms of "time-weighed" realized volatility. The price of rebalancing here is

$$P_S = E^{Q_T}\left[\int_0^T \frac{dS}{S}\right] = -\ln P(0,T) - qT + \rho\sigma^r \int_0^T \sigma^s(S,t)(T-t)dt. \tag{5.11}$$

So, from here, an increase in correlation and volatility of the short rate increases the price of rebalancing.

This is the beginning of the work of looking at variance swaps from the side of hedging and risk. Current hedges under deterministic interest rates have not performed well because the Federal Reserve has been cutting the funds with the falling equity market. This has created correlation between equities and the short rate. The frequent cuts have made the short rate more volatile. Risk must be well understood especially in such a volatile market.

Appendix A

Preliminaries

A.1 VGSSD: theory and implementation

In this dissertation work, reliable data for options is available up to a year and a half on a given day. However, there is a need to have option prices for any strike and maturity of up to three years. Perhaps a cubic spline can be used to find intermediate option values, but it is not advisable to extrapolate with a spline outside of available data. More importantly, interpolation may introduce arbitrage. Thus a calibrated arbitrage-free model that would generate option prices for any desired strike and maturity is a natural choice.

In general, one cannot fit a model with a single set of parameters across time. Recently Carr *el al.* [4] developed the theory for several four parameter self-similar processes of independent increments with self-decomposable law for the unit time distribution which adequately synthesizes European option prices across time and strikes. This section will give a review for one of these - VGSSD, which is the celebrated Variance Gamma adapted in time. Fast Fourier transform (FFT) can be utilized to quickly obtain option prices [5].

General Theory

Definition. The distribution of a random variable X is said to be self-decomposable if for any constant c, $0 < c < 1$, there exists an independent random variable $X^{(c)}$ such that $X \overset{law}{=} cX + X^{(c)}$.

The characteristic function of these laws has the form:

$$E[e^{iuX}] = \exp\left[ibu - \frac{1}{2}Au^2 + \int_{-\infty}^{\infty} (e^{iux} - 1 - iux\mathbf{1}_{|x|<1})\frac{h(x)}{|x|}dx\right],$$

where $A \geq 0$, b is a real constant, $h(x) \geq 0$, $\int_{-\infty}^{\infty} (|x|^2 \wedge 1)\frac{h(x)}{|x|}dx < \infty$, and $h(x)$ is increasing for negative x and decreasing for positive x.

A self-similar process is defined as a stochastic process $(Y(t), t \geq 0)$ with the property that for any $\lambda > 0$ and all t,

$$Y(\lambda t) \overset{law}{=} a(\lambda)Y(t).$$

Now, since

$$Y(\lambda\mu t) \overset{law}{=} a(\lambda\mu)Y(t) \overset{law}{=} a(\lambda)Y(\mu t) \overset{law}{=} a(\lambda)a(\mu)Y(t),$$

then $a(t) = t^{\gamma}$ for some γ and $Y(t)$ is γ-self-similar.

A law is self-decomposable iff it is the law at unit time of an additive process, that is also a self-similar process. This result was established by Sato [25]. Now let $Y(t)$ be the value at time t of a self-similar additive process with paths of bounded variation. Then for some time-dependent Lévy system $g(y, t)$, the characteristic function for $Y(t)$ may be written as

$$E[\, e^{iuY(t)}] = \exp\left[\int_0^t \int_{-\infty}^{\infty} (e^{iuy} - 1)g(y, s)dyds\right]. \qquad \text{(A.1)}$$

Let X be a self-decomposable random variable at unit time of some pure jump Lévy process with bounded variation, satisfying $Y(1) \overset{law}{=} X$ and

$$E[\, e^{iuX}] = \exp\left[\int_{-\infty}^{\infty} (e^{iux} - 1)\frac{h(x)}{|x|}dx\right]. \qquad \text{(A.2)}$$

Theorem. Given a self-decomposable law for the time one distribution with a characteristic function satisfying Equation A.1, then there exists a self-similar process $Y(t)$ defined with respect to the increasing scaling function t^γ and which satisfies Equation A.2 when $g(y,t)$ is equal to

$$-\frac{h'(\frac{y}{t^\gamma})\gamma}{t^{1+\gamma}},\ y > 0 \qquad and \qquad \frac{h'(\frac{y}{t^\gamma})\gamma}{t^{1+\gamma}},\ y < 0.$$

VGSSD

Define the risk-neutral process $S(t)$ in terms of $Y(t)$ as

$$S(t) = S(0)e^{rt}\frac{e^{Y(t)}}{E[e^{Y(t)}]}.$$

If $\phi_{Y(t)}(u) = E[e^{iuY(t)}]$, then the characteristic fuction of $\ln S(t)$ is

$$E[e^{iu \ln S(t)}] = e^{\left[iu\left(\ln S(0)\right)+rt-\ln\left(\phi_{Y(t)}(-i)\right)\right]}\phi_{Y(t)}(u).$$

Now, the VG process is defined as $X_{VG}(t;\sigma,\nu,\theta) = \theta G(t;\nu) + \sigma W(G(t;\nu))$ with Lévy density of the form

$$k_{VG}(x) = \begin{cases} C\frac{\exp(Gx)}{|x|}, & \text{if } x < 0 \\ C\frac{\exp(-Mx)}{x}, & \text{if } x > 0, \end{cases}$$

where

$$C = \frac{1}{\nu},$$
$$G = \left(\sqrt{\frac{\theta^2\nu^2}{4} + \frac{\sigma^2\nu}{2}} - \frac{\theta\nu}{2}\right)^{-1},$$
$$M = \left(\sqrt{\frac{\theta^2\nu^2}{4} + \frac{\sigma^2\nu}{2}} + \frac{\theta\nu}{2}\right)^{-1},$$

and $G(t;\nu)$ is a gamma process. In this case

$$h_{VG}(x) = \begin{cases} Cexp(Gx), & \text{if } x < 0 \\ Cexp(-Mx), & \text{if } x > 0. \end{cases}$$

Thus unit time VG is a self-decomposable law.

The characteristic function of $X(1)$ is given by

$$E[e^{iuX(1)}] = \left(\frac{1}{1 - iu\theta\nu + \frac{\sigma^2\nu}{2}u^2} \right)^{\frac{1}{\nu}}.$$

We want the law of $Y(t) = t^{\gamma}X(1)$ so the characteristic function for $Y(t)$ must be

$$\phi_{VGSSD}(u,t) = \left(\frac{1}{1 - iu\theta\nu t^{\gamma} + \frac{\sigma^2\nu}{2}u^2 t^{2\gamma}} \right)^{\frac{1}{\nu}}.$$

Given the log characteristic function for the stock at an arbitrary maturity, FFT can be employed to price options.

Calibration

Data for the calibration is taken on June 19, 2002 for S&P 500 index. Exclusions from data are the same as those provided in Carr *et al.* [4]. That is, out-of-the-money options were taken expiring between 30 days and 15 months. Options with strikes within 35% of the current index level, 1020, and whose price was greater than .00075 times the spot were used. On this day, after all the exclusions, there were 129 options left for calibration.

A minimization routing is run in Matlab to minimize, by changing parameters, the average absolute error between market and model prices. The calibrated parameters are $\theta = -0.16782$, $\nu = 0.65289$, $\sigma = 0.19732$, and $\gamma = 0.47708$. The average absolute error is 4.3% and larger error occurs for deeper out-of-the-money options. Calibration time was under one minute. The parameter values are consistent with those for S&P 500 provided in Carr *et al.* [4].

The VGSSD model was calibrated on the data with maturities up to 15 months. In this dissertation, the goal is to study mispricing of variance swap for longer maturities (e.g., 1, 2, 3 years). Because of the lack of data for those times, we must stylistically create it. Recall that the law of $Y(t)$ is the same as that of $t^\gamma X(1)$, where $X(1)$ is unit time VG random variable. Thus, to create the options surface that reflects longer term data, γ must be increased. In this work, γ was increased to 0.6.

VGSSD with $\theta = -0.16782$, $\nu = 0.65289$, $\sigma = 0.19732$, and $\gamma = 0.6$ will serve as a model that will create a realistic options surface. In reality, banks have this options surface available. But, due to the lack of data, there is a need to work with a stylistic, but realistic surface.

A.2 Hull-White model: general results

In the two models discussed in this work, Hull-White dynamics appear in the short rate and slightly different form in the random factor z of SR-SLV model. Some results are used in this dissertation about the solution to the Hull-White SDE and about mean and variance. These are clarified in this section.

Assume the following dynamics for Hull-White SDE:

$$dr_t = (\theta_t - kr_t)dt + \sigma^r dW_t^r. \tag{A.3}$$

By Itô's formula, the solution to equation (A.3) is

$$r_t = r_0 e^{-kt} + \int_0^t \theta e^{k(u-t)} du + \sigma^r \int_0^t e^{k(u-t)} dW_u^r \tag{A.4}$$

Because $e^{k(u-t)}$ is a deterministic function of t then $\int_0^t e^{k(u-t)} dW_u^r$ is Gaussian [21].

Thus r_t is distributed Gaussian. The mean of r_t is given by

$$E[r_t] = r_0 e^{-kt} + \int_0^t \theta e^{k(u-t)} du,$$

and variance is

$$V[r_t] = E\left[\left(\sigma^r e^{-kt} \int_0^t e^{ku} dW_u^r\right)^2\right] = \sigma^{r2} e^{-2kt} E\left[\int_0^t e^{2ku} du\right] = \sigma^{r2}\left(\frac{1-e^{-2kt}}{2k}\right),$$

where the second equality is given by the Itô isometry.

A.3 Calibration of the Hull-White model

This section outlines the calibration procedure for the Hull-White model as presented in Hull and White [17]. Note that since closed form for the discount bond price is available, a different approach to calibration can be taken. However, the proposed calibration in their paper is intuitive and easy to implement for swaptions and caps. Closed forms for a discount bond and swaption are derived in Overhaus *et al.* [23].

The Hull-White model in the dissertations is given by Equation A.3. The function θ_t is calibrated to the initial yield curve and parameters k and σ^r are used to fit the traded interest rate options. Overhaus *et al.* [23] shows that the Hull-White process can be rewritten as an Ornstein-Uhkenbeck model in the following way:

$$dx_t = -kx_t dt + \sigma_t^r dW_t^r \tag{A.5}$$

where

$$r_t = x_t + \bar{x}_t, \qquad d\bar{x}_t = (\theta_t - k\bar{x}_t)dt.$$

If $\bar{x}_0 = f(0,0)$, then $x_0 = 0$ and $E[x_t|F_0] = 0$. Here, $f(t,T)$ is the forward rate defined to be the rate, fixed at time t, for instanteneous borrowing at time T. Using variation of parameters or using the results from the previous section we get:

$$x_s = x_t\ e^{k(t-s)} + \int_t^s e^{k(u-s)}\sigma^r dW_u^r.$$

The tree is built for x_t and then $\bar{x}_t$ is used to fit the tree to the initial yield curve. Time spacing, not necessarily equal, should coincide with yield data and derivative's cashflows. Let node spacing be $\Delta x_i = \sigma^r\sqrt{3(t_i - t_{i-1})}$. If we are at time i and node $j\Delta x_i$, the tree branches to $(k-1)\Delta x_{i+1}$, $k\Delta x_{i+1}$, or $(k+1)\Delta x_{i+1}$ at time $i+1$. To ensure that branching probabilities are positive, k is chosen to be $(j\Delta x_i + E[dx])/\Delta x_{i+1}$, where

$$E[dx] \approx -j\Delta x_i(t_{i+1} - t_i)a.$$

Expression for transition probabilities are available in Hull and White [17]. Let r_{ij} be the interest rate at (i,j), Q_{ij} be the root Arrow-Debreu price, and V_{ij} be the value of a discount bond paying \$1 at time $i+1$ at node (i,j). The root price of this bond can then be written as

$$P_{i+1} = \sum_j Q_{ij}V_{ij}.$$

where

$$Q_{ij} = \sum_k p(i,j|i-1,k)\ \exp\big(-r_{i-1,k}(t_i - t_{i-1})\big)\ Q_{i-1,k}\ ,$$

$$V_{ij} = \exp\big(-r_{i,j}(t_{i+1} - t_i)\big), \quad r_{i,j} = x_{i,j} + \bar{x}_i\ .$$

Table A.2 presents a replicated numerical example of a recombining tree from Hull and White [17] using data in Table A.1.

Time to Maturity	Yield	Bond Price
1.5	5.00%	0.9277
1.6	5.10%	0.9216
2.0	5.25%	0.9003
2.5	5.30%	0.8759

Table A.1: Yield data for sample tree

$r_{ij}(\%)$				Q_{ij}			
		10.664				0.0806	
		9.048				0.0658	
		7.677	10.238			0.0064	0.0302
	11.663	6.514	7.370		0.1546	0.1024	0.2023
5.000	6.172	5.527	5.306	1.0000	0.6185	0.4098	0.4306
	3.266	4.689	3.820		0.1546	0.1024	0.2059
		3.979	2.750			0.0064	0.0313
		3.376				0.0664	
		2.864				0.0813	
$\bar{x}_0$	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	V_{ij}			
-2.9957	-2.7851	-2.8956	-2.9364			0.9582	
						0.9645	
						0.9698	0.9501
					0.9884	0.9743	0.9638
				0.9277	0.9938	0.9781	0.9738
					0.9967	0.9814	0.9811
						0.9842	0.9863
						0.9866	
						0.9886	

Table A.2: Hull-White recombining tree for the data given in table A.1.

In practice, it is not necessary to know θ_t or $\bar{x}_t$ after calibration and expressions for bond, caps and swaptions prices are available using A.5 as the underlying process. Please see Chapter 3 of Overhaus *et al.* for detailed discussion.

A cap is an interest rate derivative that provides insurance against the rate of interest on a floating-rate note rising above a certain level - cap rate. When the interest rate on the floating-rate note at some reset date exceeds the cap rate, the cap provides a payoff compensating for the difference. Thus, a cap can be viewed as a portfolio of interest rate call options at each reset date. Each option is known as a caplet.

A swaption is an option on an interest rate swap. The holder of the swaption has the right to enter into the interest rate swap at a certaing time in the future.

Examples of swaptions and caps can be found in Hull [15], Chapter 26. But, the valuation of these derivatives is done easily on a tree. For the pricing of swaptions and caps, we need $P(t,T)$ to value future swap and caplet prices. In Hull-White/Vasicek, this is given by

$$P(t,T) = A(t,T)e^{-B(t,T)R}$$

where

$$\ln A(t,T) = \ln \frac{P(0,T)}{P(0,t)} - \frac{\hat{B}(t,T)}{\hat{B}(t,t+\Delta t)} \ln \frac{P(0,t+\Delta t)}{P(0,t)}$$
$$-\frac{\sigma^2}{4k}(1-e^{-2kt})\hat{B}(t,T)[\hat{B}(t,T) - \hat{B}(t,t+\Delta t)]$$

and

$$B(t,T) = \frac{\hat{B}(t,T)}{\hat{B}(t,t+\Delta t)}\Delta t, \qquad \hat{B}(t,T) = \frac{(1-e^{k(T-t)})}{k}.$$

Pricing swaps and caplets at the root of the tree reduces to simply valuating payoffs at appropriate times on the tree via Arrow-Debreu prices, remembering that payments are made in arrears.

Market prices for these derivatives are quoted in Black volatility or price in terms of some notional amount. Data was taken for June 19, 2002 and on this day a 30-year yield curve is available (at .25 year increments), 7 caps and 12 swaptions. Calibration routine is run in Matlab to minimize the mean squared error (MSE) between model and market prices. The calibrated parameters are $k = 0.025091$ and $\sigma^r = 0.011591$. The magnitude of parameters is in line with other observed parameters in other research (e.g., Hull [17]). Calibration time was 45 seconds with MSE of 2.5%. Parameters θ_t and $\bar{x}_t$ will not be discussed in the remainder of the dissertation since the work will utilize the O-U model derived from the Hull-White model and the short rate, r_t, will not be a direct function of those parameters.

A.4 SR-SLV model: Details of PDE derivation

In chapter 3, the PDEs for the joint densities are derived for SR-LV and SR-SLV models. In the provided coefficients for those PDEs, terms such as $\dot{a}/a$ appear. This section will provide more detail as to the derivation of those coefficients. Similar transformation of variables was done in Overhaus*et al.* [23] and clarification in that case was provided by Christopher Jordinson of Deutsche Bank Quantitative Products team. Coefficients for the SR-SLV model will be discussed, since SR-LV model is just a special case of SR-SLV model with $Z = 1$.

In Chapter 3, the joint density, $\phi(x, y, z, t)$, in SR-SLV model is built from the risk-neutral processes given by:

$$dx_t = \left(g_t(y_t) - \frac{1}{2}\sigma^{s2}_{*t} e^{2z_t} \right) dt + \sigma^s_{*t} e^{z_t} dW^s_t, \tag{A.6}$$

$$dy_t = -k y_t dt + \sigma^r dW^r_t, \tag{A.7}$$

$$dz_t = \nu(\zeta_t - z_t) dt + \lambda dW^z_t, \tag{A.8}$$

where the correlation between the stock price and short rate Brownian increments is set to ρ. The PDE for $\phi(x, y, z, t)$ has a delta function for the initial distribution, thus the following transformation were proposed: $\phi(x, y, z, t) = \phi'(x', y', z', t)/(a_t b_t c_t)$, $x' = x/a_t$, $y' = y/b_t$, and $z' = z/c_t$. The variables a_t, b_t, and c_t represent standard deviations for processes x, y and z (approximate for x). Then we have (omitting subscripts):

$$\frac{\partial \phi}{\partial t} = \frac{1}{abc}\left[\frac{\partial \phi'}{\partial t} - \frac{\dot{a}}{a}\phi' - \frac{\dot{b}}{b}\phi' - \frac{\dot{c}}{c}\phi'\right],$$

where $\frac{\partial \phi'}{\partial t}$ is in terms of x and y. Expanding the derivative in t results in,

$$\frac{\partial \phi}{\partial t} = \frac{1}{abc}\left[\frac{\partial \phi'}{\partial t} + \frac{\partial \phi'}{\partial x'}\frac{\partial x'}{\partial t} + \frac{\partial \phi'}{\partial y'}\frac{\partial y'}{\partial t} + \frac{\partial \phi'}{\partial z'}\frac{\partial z'}{\partial t} - \frac{\dot{a}}{a}\phi' - \frac{\dot{b}}{b}\phi' - \frac{\dot{c}}{c}\phi'\right],$$

where $\frac{\partial \phi'}{\partial t}$ is now in terms of x' and y'. Further simplification can be made using the fact that

$$\frac{\partial x'}{\partial t} = -\frac{\dot{a}}{a}x', \qquad \frac{\partial y'}{\partial t} = -\frac{\dot{b}}{b}y', \qquad \frac{\partial z'}{\partial t} = -\frac{\dot{c}}{c}z',$$

giving

$$\frac{\partial \phi}{\partial t} = \frac{1}{abc}\left[\frac{\partial \phi'}{\partial t} - \frac{\dot{a}}{a}\left(\phi' + x'\frac{\partial \phi'}{\partial x'}\right) - \frac{\dot{b}}{b}\left(\phi' + y'\frac{\partial \phi'}{\partial y'}\right) - \frac{\dot{c}}{c}\left(\phi' + z'\frac{\partial \phi'}{\partial z'}\right)\right].$$

Noting that,

$$\frac{\partial (x'\phi')}{\partial x'} = \phi' + x'\frac{\partial \phi'}{\partial x'}$$

and applying this to terms partial involing y and z we get:

$$\frac{\partial \phi}{\partial t} = \frac{1}{abc}\left[\frac{\partial \phi'}{\partial t} - \frac{\dot{a}}{a}\frac{\partial (x'\phi')}{\partial x'} - \frac{\dot{b}}{b}\frac{\partial (y'\phi')}{\partial y'} - \frac{\dot{c}}{c}\frac{\partial (z'\phi')}{\partial z'}\right].$$

Building up the PDE, now add the first partial derivative terms to both sides:

$$\frac{\partial\phi}{\partial t}+\frac{\partial}{\partial x}\Big[\Big(g_t(y)-\frac{1}{2}\sigma_*^{s2}e^{2z}\Big)\phi\Big]-\frac{\partial}{\partial y}\Big[ky\phi\Big]+\frac{\partial}{\partial z}\Big[\nu(\zeta_t-z)\phi\Big]=\frac{1}{abc}\Bigg[\frac{\partial\phi'}{\partial t}-\frac{\dot{a}}{a}\frac{\partial(x'\phi')}{\partial x'}$$
$$+\frac{1}{a}\frac{\partial}{\partial x'}\Big[\Big(g_t(by')-\frac{1}{2}\sigma_*^{s2}e^{2cz'}\Big)\phi'\Big]-\frac{\dot{b}}{b}\frac{\partial(y'\phi')}{\partial y'}-\frac{\partial}{\partial y'}\Big[ky'\phi'\Big]-\frac{\dot{c}}{c}\frac{\partial(z'\phi')}{\partial z'}+\frac{\nu\zeta_t}{c}\frac{\partial\phi'}{\partial z'}-\frac{\partial}{\partial z'}\Big[z'\phi'\Big]\Bigg].$$

By definitions of b_t and c_t given in chapter 3, it can be seen that

$$\frac{\dot{b}}{b}=\frac{\sigma^{r2}}{2b^2}-k \qquad \frac{\dot{c}}{c}=\frac{\lambda^2}{2c^2}-\nu.$$

Substituting these, the PDEs simplifies to

$$\frac{\partial\phi}{\partial t}+\frac{\partial}{\partial x}\Big[\Big(g_t(y)-\frac{1}{2}\sigma_*^{s2}e^{2z}\Big)\phi\Big]-\frac{\partial}{\partial y}\Big[ky\phi\Big]+\frac{\partial}{\partial z}\Big[\nu(\zeta_t-z)\phi\Big]=\frac{1}{abc}\Bigg[\frac{\partial\phi'}{\partial t}-\frac{\dot{a}}{a}\frac{\partial(x'\phi')}{\partial x'}$$
$$+\frac{1}{a}\frac{\partial}{\partial x'}\Big[\Big(g_t(by')-\frac{1}{2}\sigma_*^{s2}e^{2cz'}\Big)\phi'\Big]-\frac{\sigma^{r2}}{2b^2}\frac{\partial(y'\phi')}{\partial y'}-\frac{\lambda^2}{2c^2}\frac{\partial(z'\phi')}{\partial z'}+\frac{\nu\zeta_t}{c}\frac{\partial\phi'}{\partial z'}\Bigg].$$

Remaining work to be done involves adding $g_t(y)\phi$ and second order terms to both sides. Once that is done both sides of the above equation are set to zero giving us the transformed PDE for ϕ':

$$0=\frac{\partial\phi'}{\partial t}+g_t(by')\phi-\frac{\dot{a}}{a}\frac{\partial(x'\phi')}{\partial x'}+\frac{1}{a}\frac{\partial}{\partial x'}\Big[\Big(g_t(by')-\frac{1}{2}\sigma_*^{s2}e^{2cz'}\Big)\phi'\Big]-\frac{\sigma^{r2}}{2b^2}\frac{\partial(y'\phi')}{\partial y'}-\frac{\lambda^2}{2c^2}\frac{\partial(z'\phi')}{\partial z'}$$
$$+\frac{\nu\zeta_t}{c}\frac{\partial\phi'}{\partial z'}-\frac{1}{2a^2}\frac{\partial^2(\sigma_*^{s2}e^{2cz'}\phi')}{\partial x'^2}-\frac{\sigma^{r2}}{2b^2}\frac{\partial^2\phi'}{\partial y'^2}-\frac{\lambda^2}{2c^2}\frac{\partial^2\phi'}{\partial z'^2}-\frac{\rho\sigma^r}{ab}\frac{\partial^2(\sigma_*^{s}e^{cz'}\phi')}{\partial x'\partial y'}.$$

Note that $a=\sigma_{atm}\sqrt{t}$. Since a closed form of the derivative of a does not exist here, it is approximated by a backward difference. At this stage, a little more calculus is needed to expand the derivative and isolate the partials of the densities alone. The resulting work gives rise to the form of the PDE for $\phi'(x',y',z',t)$ seen in Chapter 3:

$$0=\frac{\partial\phi'}{\partial t}+q\phi'+q1\frac{\partial\phi'}{\partial x'}+q2\frac{\partial\phi'}{\partial y'}+q3\frac{\partial\phi'}{\partial z'}+q11\frac{\partial^2\phi'}{\partial x'^2}+q22\frac{\partial^2\phi'}{\partial y'^2}+q33\frac{\partial^2\phi'}{\partial z'^2}+q12\frac{\partial^2\phi'}{\partial x'\partial y'},$$

where

$$
\begin{aligned}
q &= g_t(by') - \frac{e^{2z'c}}{2a}\frac{\partial\sigma^{s2}}{x'} - \frac{\dot{a}}{a} - \frac{\sigma^{r2}}{2b^2} - \frac{e^{2z'c}}{2a^2}\frac{\partial^2\sigma^{s2}}{\partial x'^2} - kz', \\
q1 &= \frac{g_t(by')}{a} - \frac{e^{2z'c}\sigma^{s2}}{2a} - \frac{\dot{a}x'}{a} - \frac{e^{2z'c}}{a^2}\frac{\partial\sigma^{s2}}{x'}, \\
q2 &= -\frac{\sigma^{r2}y'}{2b^2} - \frac{\rho\sigma^r e^{z'c}}{ab}\frac{\partial\sigma^s}{\partial x}, \\
q3 &= \frac{\nu\zeta}{c} - kz', \\
q11 &= -\frac{e^{2z'c}\sigma^{s2}}{2a^2}, \\
q22 &= -\frac{\sigma^{r2}}{2b^2}, \\
q33 &= -\frac{\lambda^2}{2c^2}, \\
q12 &= -\frac{\rho e^{z'c}\sigma^s\sigma^r}{ab}.
\end{aligned}
$$

The work is simpler for the SR-LV model, since it does not involve the third factor, z.

A.5 SR-SLV model: Details of second-order simulation

This section contains the calculations for the second-order method implemented for the SR-SLV model in order to reduce variance. As in Chapter 4 the setting is:

$$
\begin{aligned}
dx_t &= \left(g_t(y_t) - \frac{1}{2}\sigma_{*t}^{s2}e^{2z_t} + \rho\sigma^r\sigma_{*t}^s e^{z_t}\hat{B}(t,T)\right)dt + \sigma_{*t}^s e^{z_t}dW_t^{sQ_T}, \\
dy_t &= -\left(ky_t + \sigma^{r2}\hat{B}(t,T)\right)dt + (\sigma_1 dW_t^{rQ_T} + \sigma_2 dW_t^{rQ_T}), \\
dz_t &= \nu(\zeta_t - z_t)dt + \lambda dW_t^{zQ_T},
\end{aligned}
$$

where the Brownian motions are pairwise independent, $\sigma^{r2} = \sigma_1^2 + \sigma_2^2$, and $\sigma_1 = \sigma^r\rho$ (with ρ being the original correlation between stock and interest rate).

The second-order method from Glasserman [13] for process x_n, with $h = (t_{i+1} - t_i)$ is given as:

$$x_n(t_{i+1}) = x_n(t_i) + a_n h + \sum_{k=1}^{m} b_{nk}\Delta W^k(t_{i+1}) + \frac{1}{2}L^0 a_n h^2 + \frac{1}{2}\sum_{k=1}^{m}(L^k a_n + L^0 b_{nk})\Delta W^k(t_{i+1})$$

$$\frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{m} L^j b_{nk}(\Delta W^j(t_{i+1})\Delta W^k(t_{i+1}) - V_{jk}), \qquad \text{(A.9)}$$

where the differential operators are defined as

$$L^0 = \frac{\partial}{\partial t} + \sum_{n=1}^{d} a_n \frac{\partial}{\partial x_n} + \frac{1}{2}\sum_{n,l=1}^{d}\sum_{k=1}^{m} b_{nk} b_{lk} \frac{\partial^2}{\partial x_n \partial x_l}$$

and

$$L^k = \sum_{n=1}^{d} b_{nk} \frac{\partial}{\partial x_n},$$

for $k = 1, ...m$. Also, $V_{jk} = V_{kj}$ and is defined as a random variable taking values of $-h$ and h with probability of 1/2 each; define $V_{jj} = h$.

Then the second-order discretization for x_t or $x_1 = x$ in SR-SLV setting becomes:

$$x(t_{i+1}) = x(t_i) + a_1 h + b_{11}\Delta W^s(t_{i+1}) + \frac{1}{2}L^0 a_1 h^2 + \frac{1}{2}\sum_{k=1}^{3} L^k a_n + L^0 b_{11}\Delta W^s(t_{i+1})$$

$$\frac{1}{2}\sum_{j=1}^{3} L^j b_{11}(\Delta W^j(t_{i+1})\Delta W^s(t_{i+1}) - V_{j1}),$$

where $W^1 = W^s$, $W^2 = W^r$, and $W^3 = W^z$.

In the SR-SLV model coefficients for this method are:

$$a_1 = g_t(y_t) - \frac{1}{2}\sigma_{*t}^{s2} e^{2z_t} + \rho\sigma^r\sigma_{*t}^s e^{z_t}\hat{B}(t,T),$$

$$a_2 = -ky_t - \sigma^{r2}\hat{B}(t,T),$$

$$a_3 = \nu(\zeta_t - z_t),$$

$$b_{11} = \sigma^s_{*t} e^{z_t}, \qquad b_{12} = 0, \qquad b_{13} = 0,$$

$$b_{21} = \sigma_1, \qquad b_{22} = \sigma_2, \qquad b_{23} = 0,$$

$$b_{31} = 0, \qquad b_{23} = 0, \qquad b_{33} = \lambda.$$

Then the operators in this setting have the simplified form (dropping time subscripts):

$$\begin{aligned}
L^0 &= \frac{\partial}{\partial t} + a_1 \frac{\partial}{\partial x} + a_2 \frac{\partial}{\partial y} + a_3 \frac{\partial}{\partial z} + \frac{1}{2}\left[b_{11}^2 \frac{\partial^2}{\partial x^2} + (b_{22}^2 + b_{33}^2) \frac{\partial^2}{\partial y^2} + b_{33}^2 \frac{\partial^2}{\partial z^2} \right], \\
L^1 &= b_{11} \frac{\partial}{\partial x} + b_{21} \frac{\partial}{\partial y}, \\
L^2 &= b_{22} \frac{\partial}{\partial y} \\
L^3 &= b_{33} \frac{\partial}{\partial z}.
\end{aligned}$$

Let $\hat{B}(t,T) = (1 - \exp(-kh))/k = \hat{B}$, $g_t(y_t) = g$, and $\sigma^s_{*t} = \sigma^s_*$. All quantities on the right hand side will be evaluated at time t_i and $S(t_i)$ where appropriate. Thus, for simplicity, time and space subscripts will be dropped. Substituting these into equation (A.10) for the process x_t (A.7) yields:

$$\begin{aligned}
x(t_{i+1}) = x(t_i) &+ a_1 h + b_{11} \Delta W_1 + \frac{h^2}{2}\Bigg[-\rho\sigma^r \sigma^s_* e^{-k(T-t_i)} + a_1\Big(-\frac{1}{2} e^{2z} \frac{\partial \sigma^{s2}_*}{\partial x} + \rho\sigma^r \frac{\partial \sigma^s_*}{\partial x} e^z \hat{B} \Big) \\
&+ a_2 \frac{\partial g}{\partial y} + a_3 \Big(\rho\sigma^r \sigma^s_* e^z \hat{B} - (\sigma^s_* e^z)^2 \Big) + \frac{1}{2}\Bigg(b_{11}^2 \Big(-\frac{1}{2} e^{2z} \frac{\partial^2 \sigma^{s2}_*}{\partial x^2} + \rho\sigma^r \frac{\partial^2 \sigma^s_*}{\partial x^2} e^z \hat{B} \Big) + (\sigma_1 + \sigma_2) \frac{\partial^2 g}{\partial y^2} \Bigg) \\
&+ \lambda^2 \Big(-2(\sigma^s_* e^z)^2 + \rho\sigma^r \sigma^s_* e^z \hat{B} \Big) \Bigg] + \frac{h}{2}\Bigg[b_{11} \Big(-\frac{1}{2} e^{2z} \frac{\partial \sigma^{s2}_*}{\partial x} + \rho\sigma^r \frac{\partial \sigma^s_*}{\partial x} e^z \hat{B} \Big) \Delta W_1 + \sigma_1 \frac{\partial g}{\partial y} \Delta W_1 \\
&+ \sigma_2 \frac{\partial g}{\partial y} \Delta W_2 + \lambda \Big(-(\sigma^s_* e^z)^2 + \rho\sigma^r \sigma^s_* e^z \hat{B} \Big) \Delta W_3 + \Big[a_1 e^z \frac{\partial \sigma^s_*}{\partial x} + a_3 \sigma^s_* e^z + \frac{1}{2}\Big(b_{11}^2 e^z \frac{\partial^2 \sigma^s_*}{\partial x^2} \\
&+ \lambda^2 \sigma^s_* e^z \Big) \Big] \Delta W_1 \Bigg] + \frac{1}{2}\Bigg[\sigma^s_* e^{2z} \frac{\partial \sigma^s_*}{\partial x} (\Delta W_1^2 - h) + \lambda \sigma^s_* e^z (\Delta W_1 \Delta W_3 + \xi) \Bigg],
\end{aligned}$$

where ξ is a random variable taking values of $-h$ and h with probability .5 and independent of the Brownian motions. And, the Brownian motions are independent from each other.

Remembering that from $y_0 = 0$ we have

$$g_t(y_t) = r_t - f(0,t) = y_t + \sigma^r \int_0^t \hat{B}(s,t)\exp(-\Lambda_{st})ds.$$

Then it is true that:

$$\frac{\partial g}{\partial y} = 1, \qquad \frac{\partial^2 g}{\partial y^2} = 0, \qquad \frac{\partial g}{\partial t} = 0.$$

After much simplifying, the discretization can be written as

$$\begin{aligned} x(t_{i+1}) = x(t_i) &+ A\Delta W_1 + B\Delta W_2 + C\Delta W_3 + D \\ &+ \frac{b_{11}}{2}\Big[e^z \frac{\partial \sigma_*^s}{\partial x}(\Delta W_1^2 - h) + b_{33}(\Delta W_3 \Delta W_1 + \xi)\Big], \end{aligned}$$

where

$$\begin{aligned} A &= \frac{h}{2}\Big[b_{11}\Big(-\frac{e^{2z}}{2}\frac{\partial \sigma_*^{s2}}{\partial x} + \rho\sigma^r \frac{\partial \sigma_*^s}{\partial x} e^z \hat{B}\Big) \\ &\quad + \Big[a_1 e^z \frac{\partial \sigma_*^s}{\partial x} + a_3 \sigma_*^s e^z + \frac{1}{2}\Big(b_{11}^2 e^z \frac{\partial^2 \sigma_*^s}{\partial x^2} + \lambda^2 \sigma_*^s e^z\Big)\Big]\Big] + b_{11} + \frac{h}{2} b_{21}, \\ B &= b_{22}, \\ C &= \lambda\Big(-b_{11}^2 + \rho\sigma^r \sigma_*^s e^z \hat{B}\Big), \\ D &= a_1 h + \frac{h^2}{2}\Big[a_1\Big(-\frac{e^{2z}}{2}\frac{\partial \sigma_*^{s2}}{\partial x} + \rho\sigma^r \frac{\partial \sigma_*^s}{\partial x} e^z \hat{B}\Big) + a_2 + a_3\Big(-b_{11}^2 + \rho\sigma^r \sigma_*^s e^z \hat{B}\Big) \\ &\quad - \rho\sigma^r \sigma_*^s e^{-kh} + \frac{1}{2}\Big(b_{11}^2\Big(-\frac{e^{2z}}{2}\frac{\partial^2 \sigma_*^{s2}}{\partial x^2} + \rho\sigma^r \frac{\partial^2 \sigma_*^s}{\partial x^2} e^z \hat{B}\Big) \\ &\quad + \lambda^2\Big(-2b_{11}^2 + \rho\sigma^r \sigma_*^s e^z \hat{B}\Big)\Big]. \end{aligned}$$

To simplify things, we can let

$$\frac{\partial \sigma_*^{s2}}{\partial x} = 2\sigma_*^s \frac{\partial \sigma_*^s}{\partial x},$$

and

$$\frac{\partial^2 \sigma_*^{s2}}{\partial x^2} = 2\left(\frac{\partial \sigma_*^s}{\partial x} + \sigma_*^s \frac{\partial^2 \sigma_*^s}{\partial x^2}\right).$$

Here ξ is a random variable independent of the Brownian motions taking values of $-h$ and h with probability of .5 each.

This method required the knowledge of the first two derivatives of local volatility with respect to x. Central differences can be used to approximate these after fitting a spline over the local volatility curve at any given time. Care must be taken on and outside of the boundaries where local volatility is defined. Local volatility is about twice as steep as implied volatility and using a spline to extrapolate local volatility outside of the defined region is not desirable. Very large and unrealistic values of volatility can be obtained this way. Thus, derivatives and values of volatility outside of the boundaries are defined as values on the boundary.

BIBLIOGRAPHY

[1] Atlan, M. Localizing Volatilities. Working Paper, Laboratorie de Probabilitiés, Université Pierre at Marie Curie, (2006).

[2] Black, F. and P. Karasinski. Bond and Option Pricing When Short Rates are Lognormal. *Financial Analysts Journal* Jul/August: 52-59, (1991).

[3] Cont, R. and P. Tankov. Financial Modelling with Jump Processes. Chapman & Hall, (2004).

[4] Carr, P., H. Geman, D. Madan, and M. Yor. Self-Decomposability and Option Pricing. *Mathematical Finance* 17(1): 31-57,(2007).

[5] Carr, P. and D. Madan. Option valuation using the FFT. *Journal Computational Finance* 2: 61-73, (1999).

[6] Craig, I., and A. Sneyd. An Alternating-Direction Implicit Scheme for Parabolic Equations with Mixed Derivatives. *Computational Mathematical Applications* 16(4): 341-350, (1988).

[7] Derman, E.,*et al.* More Than You Ever Wanted To Know About Volatility Swaps. Goldman Sachs: Quantitative Strategies Research Notes, (1999).

[8] Derman, E., I. Kani, and J. Zou. The Local Volatility Surface. Goldman Sachs: Quantitative Strategies Research Notes, (1995).

[9] Douglas, J. and H. Rachford. On the numerical solution of heat conduction problems in two and three space variables. *Trans. American Mathematical Society* 82:421-439, (1956).

[10] Duffy, D. Finite Difference Methods in Financial Engineering: A Partial Differential Equation Approach. Wiley Finance, (2006).

[11] Dupire, B. Pricing with a Smile. *Risk Magazine* 7: 18-20,(1994).

[12] Gatheral, J. The Volatility Surface: A Practitioner's Guide. Wiley Finance, (2006).

[13] Glasserman, P. Monte Carlo Mathods in Financial Engineering. Springer, (2004).

[14] Gyöngy, I., Mimicking the One-Dimensional Marginal Distributions of Processes Having an Itô Differential, *Probability Theory and Related Fields* 71: 501-516, (1986).

[15] Hull, J. Options, Futures, and Other Derivatives (6th edition). Prentice Hall, (2006).

[16] Hull, J., and A. White. Pricing Interest Rate Derivative Securities. *Review of Financial Studies* 3(4): 573-592, (1990).

[17] Hull, J., and A. White. The General Hull-White Model and Super Calibration. *Financial Analysts Journal* 57(6): 34-43, (2001).

[18] Li, M., and N. Pearson. Price Deviations of S&P 500 Index Options from the Black-Scholes Formula Follow a Simple Pattern. University of Illinois at Urbana-Champaign: Working Paper, (2004).

[19] Madan, D., M. Qian, and Y. Ren. Calibrating and Pricing with Embedded Local Volatility Models. *Risk Magazine* 9: 138-143, (2007).

[20] Madan, D. and E. Seneta. The variance gamma (VG) model for share market returns. *Journal of Business* 63:511-524,(1990).

[21] Mamon, R. Three ways to Solve for Bond Prices in the Vasicek Model. *Journal of Applied Mathematics and Decision Sciences* 8(1): 1-14, (2004).

[22] McKee, S., D. Wall, and S. Wilson. An Alternating Direction Implicit Scheme for Parabolic Equations with Mixed Derivatives and Convective Terms. *Journal of Computational Physics* 126: 64-76, (1996).

[23] Overhaus, M., *et al*. Equity Hybrid Derivatives. Wiley Finance, (2007).

[24] Jackwerth,J and M. Rubinstein. Recovering Probability Distributions from Option Prices. *Journal of Finance* 51(5): 1611-1631, (1996).

[25] Sato, L. Self-similar Processes with independent Increments. *Probability Theory Related Fields* 89: 285-300, (1991).

[26] Schoutens, W. Lévy Processes in Finance. Wiley, (2005).

[27] Schoutens, W. Moment Swaps. Leuven Statistics Research Centre Report, (2005).

[28] Tuckman, B. Fixed Income Securities (2nd edition). Wiley Finance, (2002).

[29] Vasicek, O. An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics* 5(2): 177-188, (1977).

[30] Wilmott, Paul. Derivatives: The Theory and Practice of Financial Engineering. Wiley, (1998).

[31] Yanenko, N. The Method of Fractional Steps. Springer-Verlang, Berlin (1971).