# Option Pricing Models with Stochastic Volatility and Jumps




Farhan Kalsheker

12 October 2009

A Dissertation submitted in partial fulfillment for the degree of Master of Science
Department of Statistical Sciences and Department of Mathematics and Applied Mathematics
University of Cape Town, South Africa

Supervisor: Professor Renkuan Guo

# Declaration

I, Farhan Kalsheker, hereby,

1. grant the University of Cape Town free licence to reproduce this dissertation in whole or in part, for the purpose of research;

2. declare that:

    i. this dissertation is my own unaided work, both in concept and execution,

    ii. each significant contribution to this dissertation from the work of other people has been cited and referenced.

    iii. I know the meaning of Plagiarism and declare that all of the work in the document, save for that which is properly acknowledged, is my own

    iv. neither the substance nor any part of this dissertation has been submitted in the past, or is being, or is to be submitted for a degree at this University or at any other university.

I am now presenting this dissertation for examination for the degree of Master of Science.



_____

Signature

_____

Date

# Abstract

Exotic equity options are specialized instruments which are typically traded over the counter. Their prices are primarily determined by option pricing models which should be able to price exotic options consistently with the market prices of corresponding vanilla options. Additionally, option pricing models should have intuitive dynamics which are able to capture real world behavior (such as stochastic volatility effects and jumps in the price of the underlying).

This dissertation tackles the question of which option pricing model to use; it compares diffusion, pure jump and jump-diffusion models. All models are fitted to one-day price data on S&P500 European vanilla options; the models with the best fit exhibit the smallest error in pricing between model prices and market prices.

The stochastic volatility with jumps (SVJ) models are found to perform the best. The SVJ-DE model, a new variant of this type of model (which is based on Heston-type stochastic volatility and Kou-type double exponential jumps in the log price), is presented and tested. The Heston SV model is ranked third best. There is a significant performance gap between the SV/SVJ models and the remaining models. The variance-gamma model with stochastic time is found to be the best performing model from the pure jump and simple jump-diffusion categories. The Kou jump-diffusion model with double exponential jumps and constant diffusion volatility ranks next, followed by the Merton jump-diffusion model and the variance-gamma pure jump model.

On comparison of model and market implied volatility surfaces, the pure jump and simple jump-diffusion models are found to be efficient at generating volatility smile effects, but not volatility skew effects. The converse holds for the Heston SV model. The SVJ models exploit this behavior in an attempt to use the jump component to generate the smile effects on the short end of the volatility surface and the stochastic volatility diffusion component to generate the skew effects on the long end of the volatility surface.

The application of the SV and SVJ models is demonstrated by computing the prices of barrier options via Monte Carlo simulation. Both of the SVJ models give similar barrier option prices.

Diffusion processes and jump processes are the two main building blocks of any option pricing model. This research finds that simple jump-diffusion models and pure jump models are unable to demonstrate good performance when fitting to a complete grid of market option prices. The Heston stochastic volatility pure diffusion model gives better performance compared to these jump models.

The SVJ models which have both a stochastic volatility diffusion component and a jump component are found to give the best performance. The SVJ-DE model has the added advantage of being able to generate upward and downward jumps from different exponential distributions, versus the Bates model which generates jumps from a normal distribution.

# Acknowledgements

My sincere thanks to my supervisor, Professor Renkuan Guo, for his invaluable advice and directions, and for his continuous support and help in my dissertation.

Special thanks to my parents (and brothers) for their support, guidance and prayers.

Thanks to Aidan for his advice.

A final thanks to Dr Danni Guo for providing the dissertation template.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction



European vanilla options on thousands of underlying assets (including stocks and indices) are commonly exchange traded, standardized and can be very liquid; this results in the price of these options being determined primarily by market forces, including supply and demand as well as market sentiment. In contrast to vanilla options, exotic options are primarily traded over-the-counter (OTC), are not usually standardized nor are they liquid; therefore, the prices of these options are not quoted on securities exchanges. It becomes necessary to price these options using an option pricing model, and the model should price the exotic options consistently with the prices of the corresponding vanilla options.

The Black-Scholes model is currently primarily used to quote option prices in the form of implied volatilities. Implied volatility surfaces formed from option market price grids show both volatility skew (slope) and smile (curvature) effects, see Cont and Fonseca (2002).

The Black-Scholes model cannot be used to accurately price vanilla options (hence exotic options), and the fact that the model assumes constant volatility of the underlying process is arguably one of the biggest problems with the model. A study on the empirical properties of underlying asset processes (including the S&P500 index) can be found in Cont (2001) and Cont and Tankov (2003). The latter source studies the nature of asset price movements, noting that empirically, asset prices move by jumps and are not continuous. This implies that pure diffusion models have counter-intuitive dynamics; furthermore, this implies that models which are to be intuitively correct representations of reality should be purely jump driven or should at least have a jump component.

It should be noted that the local volatility models of Dupire (1994), Derman and Kani (1998) and Derman et al. (1996) became a popular class of models which were able to give arbitrage-free prices whilst capturing volatility skew and smile effects. Option prices are obtained from these models using tree methods. However, Hagan et al. (2002) finds that the dynamics of the volatility smile under these models are not consistent with market dynamics.

This dissertation will focus on stochastic volatility (SV) diffusion models, jump models and stochastic volatility plus jump (SVJ) models.

Chapter 2 gives an introduction into risk-neutral pricing and applies this theory when the volatility of the underlying asset is assumed to follow a stochastic process; the Heston (1993) model is covered. The Heston model was chosen for numerical implementation and testing because it has

been one of the most successful SV models. Other popular stochastic volatility models are the Hull and White (1987) and Hagan et al. (2002) models.

Chapter 3 develops the theory of Lévy processes and option pricing under Lévy processes (including option pricing via the characteristic function). Lévy processes are fundamental to the construction of the option pricing models with jumps which follow in chapter 4. The jump-diffusion models of Merton (1976) and Kou (2002) are covered, as well as the pure jump variance-gamma (VG) model and the VG model with Cox-Ingersoll-Ross stochastic time (VG-CIR).

Chapter 5 introduces the Bates (1996) jump-diffusion model which incorporates both Heston-type stochastic volatility and jumps (normal jumps in the log-asset price). This chapter also introduces a jump-diffusion model which combines Heston-type stochastic volatility with Kou-type jumps (double-exponential jumps in the log-asset price).

Chapter 6 gives details on the market data used for numerical comparisons, the market implied volatility surface (including the method for obtaining a smoothed version of the market implied volatility surface). The chapter also gives details on model calibration, where each model is fitted to a grid of option prices on the S&P500 index. This is essentially how each models parameters are calculated so as to ensure that the model is pricing consistently with market prices, details are also given on the optimization method chosen to perform the calibration (whilst both global and local optimization methods are used, the ASA global optimization method is the primary method used).

Chapter 7 compares models based on their ability to price consistently with market prices. Comparison is based on implied volatility surfaces of the models which gives an indication of how model implied volatilities compare to market implied volatilities. Finally, the prices of barrier options are calculated via Monte Carlo simulation (under the best performing models). The purpose of this comparison is to determine, from a numerical and practical point of view, which type of option pricing model to use for the pricing of exotic equity options.

# Chapter 2

# Stochastic Volatility Models

## 2.1 Risk-Neutral Pricing

This section gives a brief overview of risk-neutral pricing, more detailed information can be found in sources such as Bingham and Kiesel (2004), Bjork (2004) and Hunt and Kennedy (2004). Assume we are working on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$. There are *d+1* traded assets, whose price processes are given by stochastic processes $S^0, \ldots, S^d$ and where $S^0$ is a numéraire which is taken to be the risk-free asset. The discounted asset price is defined as $\tilde{S}^i_t = S^i_t/S^0_t, i = 1, 2, \ldots d$.

**Definition 2.1** (Arbitrage Opportunity)**.** (Bingham and Kiesel, 2004, §6.1.3) A self-financing trading strategy $\varphi$ is called an *arbitrage opportunity* if the value of the portfolio $V(\varphi)$ satisfies the following set of conditions:

- $V_0(\varphi) = 0$
- $\mathbb{P}(V_T(\varphi) \geq 0) = 1$
- $\mathbb{P}(V_T(\varphi) > 0) > 0$

where the value of the portfolio $\varphi$ at time $t$ is given by the scalar product $V_t(\varphi) := \varphi_t \cdot S_t = \sum_{i=0}^d \varphi^i_t S^i_t,\ t \in [0, T]$, the gains process $G_t(\varphi)$ is defined by $G_t(\varphi) := \int_0^t \varphi_u dS_u = \sum_{i=0}^d \int_0^t \varphi^i_u dS^i_u$, and a trading strategy $\varphi$ is called *self-financing* if the wealth process $V_t(\varphi)$ satisfies $V_t(\varphi) = V_0(\varphi) + G_t(\varphi)$ for all $t \in [0, T]$

**Definition 2.2.** A probability measure $\mathbb{Q}$ defined on $(\Omega, \mathcal{F})$ is an equivalent (local) martingale measure (EMM) if:

- $\mathbb{Q} \sim \mathbb{P}$ (i.e. $\mathbb{Q}$ is equivalent to $\mathbb{P}$)
- the discounted price process $\tilde{S}$ is a $\mathbb{Q}$ (local) martingale.

**Theorem 2.3** (**The First Fundamental Theorem**)**.** *(Bjork, 2004, §10.5)*
*The model is arbitrage free if and only if there exists an equivalent (local) martingale measure* $\mathbb{Q}$.

Bjork (2004), Delbaen and Schachermayer (1994) and Delbaen and Schachermayer (1998) can be referred to for technical details, sufficient conditions and a proof of the above theorem. Now, if an equivalent martingale measure can be found (the model is arbitrage free), then the following theorem can be applied to give the arbitrage free price process of a contingent claim:

**Theorem 2.4** (**General Pricing Formula**). *(Bjork, 2004, §10.18)*
*The arbitrage free price process for the contingent claim X is given by*

$$\Pi_t(X) = S_t^0 \mathbb{E}^{\mathbb{Q}} \left[ \frac{X}{S_T^0} \middle| \mathcal{F}_t \right] \tag{2.1}$$

*where $\mathbb{Q}$ is the (not necessarily unique) equivalent martingale measure.*

When $S^0$ is the risk-free asset and $\mathbb{Q}$ is the risk-neutral measure (under which the discounted asset price processes have zero drift), then this is termed the risk-neutral valuation formula.

**Theorem 2.5** (**The Second Fundamental Theorem**). *(Bjork, 2004, §10.17)*
*Assume that the market is arbitrage free. Then the market is complete if and only if the equivalent martingale measure is unique.*

From the Second Fundamental Theorem, we note that in a complete market, the no arbitrage requirement will ensure a unique price for a derivative (since the equivalent martingale measure is unique). Conversely, in an incomplete market, under the no arbitrage requirement there may be several equivalent martingale measure's which lead to several arbitrage free prices for a derivative. Hence the no arbitrage requirement is insufficient to yield a unique price in an incomplete market and it will be necessary to choose the equivalent martingale measure under which to price. One can assume that the equivalent martingale measure is determined by the market, and that the model will present arbitrage free prices which are consistent with the market following calibration of model parameters, this is the approach adopted in chapter 6.

Finding and constructing the equivalent martingale measure relies on Girsanov's theorem which gives asset price dynamics under a change of measure. The theorem facilitates construction of the risk-neutral probability measure under which the drift of the discounted asset price process is zero. The theorem is applied in the next section (2.2) where volatility is stochastic and there are two sources of randomness.

**Theorem 2.6** (**Girsanov's theorem**). *(Shreve, 2004, §5.4.1)*
*Assume that $\mathbf{W}_t = (W_t^1, \ldots, W_t^d)$ is a multi-dimensional Brownian motion on probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which is equipped with filtration $(\mathcal{F}_t)$. Let the fixed final time be $T$. Let $\mathbf{\Theta}_t = (\theta_t^1, \ldots, \theta_t^d)$ be a $d$-dimensional adapted process. Define*

$$\mathbf{Z}_t = \exp\left( -\int_0^t \mathbf{\Theta}_u \cdot d\mathbf{W}_u - \frac{1}{2}\int_0^t ||\mathbf{\Theta}_u||^2 \, du \right), \tag{2.2}$$

$$\widetilde{\mathbf{W}}_t = \mathbf{W}_t + \int_0^t \mathbf{\Theta}_u \, du \tag{2.3}$$

*and assume that*

$$\mathbb{E}\int_0^T ||\mathbf{\Theta}_u||^2 \, \mathbf{Z}_u^2 \, du < \infty. \tag{2.4}$$

*Set* $\mathbf{Z} = \mathbf{Z}_T$. *Then* $\mathbb{E}\mathbf{Z} = 1$, *and under the probability measure* $\mathbb{Q}$ *given by*

$$\mathbb{Q}(A) = \int_A \mathbf{Z}(\omega) d\mathbb{P}(\omega) \quad \textit{for all} A \in \mathcal{F}_T, \tag{2.5}$$

*The process* $\widetilde{\mathbf{W}}_t$ *is a* $d$*-dimensional Brownian motion.*

where the Ito integral is defined:

$$\int_0^t \mathbf{\Theta}_u \cdot d\mathbf{W}_u = \int_0^t \sum_{j=1}^d \theta_u^j \, dW_u^j = \sum_{j=1}^d \int_0^t \theta_u^j \, dW_u^j \tag{2.6}$$

and the Euclidean norm is defined:

$$||\mathbf{\Theta}_u|| = \left( \sum_{j=1}^d (\theta_u^j)^2 \right)^{\frac{1}{2}} \tag{2.7}$$

and the multi-dimensional Brownian motion $\widetilde{\mathbf{W}}_t$ has components given by:

$$\widetilde{\mathbf{W}}_t^j = \mathbf{W}_t^j + \int_0^t \theta_u^j \, du, \quad j = 1, \ldots, d. \tag{2.8}$$

## 2.2 Pricing with Stochastic Volatility

It is widely accepted that the constant volatility assumption of the Black-Scholes model is unrealistic (for example, see Cont (2001) for an empirical study). This led to the development of stochastic volatility models which model the volatility of the underlying process with a stochastic process. As it turns out, the inclusion of stochastic volatility into an option pricing model enables the model to produce option prices which capture skew (and smile) effects; this is demonstrated in chapter 7. Note that the volatility process should ensure that only positive volatilities are allowed; a negative volatility has neither a mathematical nor a financial meaning. It is convention to include mean-reversion in the volatility process dynamics (as is used in the Heston and Bates models). This section (which follows Lee (1999) closely) aims to give an introduction to pricing when volatility is modelled as a stochastic process.

As before, assume that we are working on filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$. The dynamics of the asset price $(S_t)$ and the squared volatility process $(v_t)$ are given by:

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t^S \tag{2.9}$$

$$dv_t = \alpha_t dt + \beta_t dW_t^v \tag{2.10}$$

$$\sigma_t = \sqrt{v_t} \tag{2.11}$$

where $W^S$ and $W^v$ are $(\mathcal{F}_t)$-Brownian motions with correlation $\rho$. $W^v$ can be written in the form:

$$W_t^v = \rho W_t^S + \sqrt{1-\rho^2} W_t^{(2)}. \tag{2.12}$$

where $W_t^{(2)}$ is a Brownian motion which is uncorrelated with $W_t^S$. Assume that the market model is arbitrage-free and also assume any other stronger conditions which are necessary for the fundamental theorem of asset pricing to hold (theorem 2.3, see Delbaen and Schachermayer (1994) and Delbaen and Schachermayer (1998)), then there exists an equivalent local martingale measure $\mathbb{Q} \sim \mathbb{P}$ such that the discounted prices of all assets are local martingales under $\mathbb{Q}$.

By Girsanov's theorem, there exist $(\mathcal{F}_t)$ adapted processes $\lambda^S$ and $\lambda^{(2)}$ such that

$$\mathbb{E}_t\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right) = \exp\left( -\int_0^t \lambda_u^S dW_u^S - \int_0^t \lambda_u^{(2)} dW_u^{(2)} - \frac{1}{2}\int_0^t (\lambda_u^S)^2 + (\lambda_u^{(2)})^2 du \right) \tag{2.13}$$

and $\tilde{W}^S$ and $\tilde{W}^{(2)}$ are $((\mathcal{F}_t), \mathbb{Q})$ Brownian motions:

$$\tilde{W}_t^S := W_t^S + \int_0^t \lambda_u^S du \tag{2.14}$$

$$\tilde{W}_t^{(2)} := W_t^{(2)} + \int_0^t \lambda_u^{(2)} du \tag{2.15}$$

In order to ensure that $S$ is a martingale under $\mathbb{Q}$ we require:

$$\lambda_t^S = \frac{\mu_t - r}{\sigma_t} \tag{2.16}$$

which gives the $\mathbb{Q}$-dynamics of $S$:

$$dS_t = rS_t dt + \sigma_t S_t d\tilde{W}_t^S \tag{2.17}$$

$\lambda^S$ is termed the *asset risk premium* and $\lambda^{(2)}$ is called the *"volatility risk premium"* or the $W^{(2)}$-*risk premium*. The dynamics of $S$ do not lead to a unique choice of $\lambda^{(2)}$, i.e. different choices of $\lambda^{(2)}$ lead to different local equivalent martingale measures and different pricing functions under $\mathbb{Q}$. Since the local equivalent martingale measure is not unique, the market is incomplete. There are two sources of randomness, but only one tradeable asset $S$. So, whilst the Brownian motion of $S$ can be hedged, the Brownian motion of $v$ cannot be hedged since $v$ is not a tradeable asset. It should be noted that growth in financial markets has led to an increase in the availability and use of volatility derivatives (listed and over the counter; futures, options and swaps) to the extent that it may be possible to hedge volatility risk in certain cases (by trading these volatility derivatives). However, these instruments will not be available in *all* cases, hence we proceed with the assumption that it is not possible to hedge the Brownian motion of the volatility process, therefore the market is incomplete.

Under $\mathbb{Q}$, we have the following:

$$dS_t = rS_t dt + \sigma_t S_t d\tilde{W}_t^S \tag{2.18}$$

$$dv_t = (\alpha_t - \lambda_t^v \beta_t)dt + \rho\beta_t d\tilde{W}_t^S + \sqrt{1-\rho^2}\beta_t d\tilde{W}_t^{(2)} \tag{2.19}$$

where

$$\lambda_t^v = \rho\lambda_t^S + \sqrt{1-\rho^2}\lambda_t^{(2)}. \tag{2.20}$$

$\lambda^v$ is termed the $W^v$*-risk premium*, and is a different version of the *"volatility risk premium"*. Multiplying an options exposure to $W^{(2)}$-risk by $\lambda^{(2)}$ gives the excess drift associated with $W^{(2)}$-risk. Similarly, multiplying an options exposure to $W^v$-risk by $\lambda^v$ gives the excess drift associated with $W^v$-risk.

The equation for $S$ is in risk-neutral form, and the drift of $v$ has an arbitrary function $\lambda_t$ added (which cannot be hedged). Then set $\lambda_t = 0$, which implies that we are in the risk-neutral measure. This is the convention followed in Gatheral (2006), who assumes that the equations for $S$ and $v$ are in risk-neutral terms and that the risk-neutral measure is generated by calibrating the model to market prices, as done in chapter 6.

The risk-neutral pricing formula can then be applied to give an arbitrage free price process $\Pi_t$ for a contingent claim, with payoff function $X$, as the risk-neutral expectation:

$$\Pi_t(X) = e^{rt}\,\mathbb{E}^{\mathbb{Q}}\left[\frac{X}{e^{rT}}\,\middle|\,\mathcal{F}_t\right] \tag{2.21}$$

## 2.3 The Heston (1993) Diffusion Model

The Heston (1993) model assumes that the asset price process follows the diffusion (risk-neutral dynamics)

$$dS_t = r_t S_t dt + \sqrt{v_t} S_t dW_t^S \tag{2.22}$$

The variance process, $dv_t$, is given by the following mean reverting square root process (as in Cox et al. (1985))

$$dv_t = \kappa(\theta - v_t)dt + \sigma_v\sqrt{v_t}dW_t^v \tag{2.23}$$

where $W_t^S$ and $W_t^v$ are Brownian motion processes with correlation $\rho$. Note that if the Feller condition $2\kappa\theta - \sigma_v^2 > 0$ is satisfied, then the variance process will always be positive. The following table gives a description of the model parameters.

| | |
|---|---|
| $\kappa$ | speed of mean reversion |
| $\theta$ | long variance |
| $\rho$ | correlation |
| $\sigma_v$ | volatility of volatility |
| $v_0$ | short variance |

**Characteristic Function and Solution for European Option Prices**

The Heston (1993) model presents a solution for European vanilla option prices via the characteristic function. This requires a simple numerical integration and the fact that the model yields these solutions was arguably the most important factor which contributed to the success of this model. This is due to the fact that computation via the characteristic function is not computationally expensive when compared to computation via alternative methods, such as Monte Carlo methods.

The version of the Heston characteristic function, as used in Schoutens et al. (2003), Gatheral (2006) and Albrecher et al. (2006), is presented below in equation (2.24). This is then used for the calculation of option prices via the integral in equation 3.47:

$$\begin{aligned}\varphi_{X_T} &= \exp\left\{\frac{\kappa\theta}{\sigma^2}\left\{(\kappa-\rho\sigma iu-d)T-2\log\left[\frac{1-ge^{-dT}}{1-g}\right]\right\}\right\} \\ &\quad\times\exp\left\{\frac{v}{\sigma^2}\big(\kappa-\rho\sigma iu-d\big)\left[\frac{1-e^{-dT}}{1-ge^{-dT}}\right]\right\} \qquad (2.24)\end{aligned}$$

where

$$d = \sqrt{\big(\rho\sigma iu-\kappa\big)^2+\sigma^2\big(iu+u^2\big)} \qquad (2.25)$$

$$g = \frac{\kappa-\rho\sigma iu-d}{\kappa-\rho\sigma iu+d} \qquad (2.26)$$

It should be noted that there is a second version of the Heston characteristic function which was derived in the original paper of Heston (1993), and this second version can lead to incorrect model prices for longer maturities. The numerical instability is due to the branch cut of the complex logarithmic function which forms part of the Heston characteristic function. This problem can be remedied by using the version of the characteristic function given above. The reader is referred to Albrecher et al. (2006) for a full discussion on this topic, proof of the stability of this version, and details on the threshold maturity from which the second version exhibits numerical instability.



**Discretization**

The discretization of the asset price is given by:

$$S_{i+1} = S_i+(r-q)S_i\Delta t+\sqrt{v_i}S_i\sqrt{\Delta t}Z^S \qquad (2.27)$$

The Euler discretization of the variance process is given by

$$v_{i+1} = v_i+\kappa(\theta-v_i)\Delta t+\sigma_v\sqrt{v_i}\sqrt{\Delta t}Z^v \qquad (2.28)$$

$Z^S$ and $Z$ are independent standard normal random variables and $Z^v=\rho Z^S+\sqrt{1-\rho^2}Z$. Negative variances can result from this method, the common way to deal with this is to include a check

to either set a negative variance to zero (absorbing variance barrier) or to take its absolute value (reflecting). The negative variance problem is reduced by using a Milstein discretization scheme:

$$v_{i+1} = \left(\sqrt{v_i} + \frac{\sigma_v}{2}\sqrt{\Delta t}Z^v\right)^2 + \kappa(\theta - v_i)\Delta t - \frac{\sigma_v^2}{4}\Delta t \tag{2.29}$$

There has been significant research into the simulation of the Heston model (and other stochastic volatility models). Alternative simulation approaches include simulating the log-variance to deal with the occurrence of negative variances, using implicit discretization schemes, and exact simulation methods. More information can be found in, for example, Lord et al. (2008), Kahl and Jackel (2006) and Broadie and Kaya (2004).

# Chapter 3

# Levy Processes

## 3.1 Definition and Properties

Assume the stochastic base $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ with the filtration $(\mathcal{F}_t)_{t\geq 0}$,

**Definition 3.1 (Lévy process).** An $\mathbb{R}^d$ valued stochastic process $X = (X_t)_{t\geq 0}$ is called a Lévy process if and only if it possesses the following properties:

(i) *Stationary increments:* for $h \geq 0$, $X_{t+h} - X_t =^d X_h$.

(ii) *Independent increments:* for $h \geq 0$, $X_{t+h} - X_t$ is independent of $\mathcal{F}_t$.

(iii) *Stochastic continuity:* $\forall \epsilon \geq 0, \lim_{h\to 0}\mathbb{P}(|X_{t+h} - X_t| \geq \epsilon) = 0$, implying, $X_s \to X_t$ in probability as $s \to t$.

(iv) *Càdlàg:* Sample paths of *X* are right-continuous with left-limits.

A process which satisfies the first three conditions but not the last condition is termed a Lévy process in law. It can be shown that a Lévy process in law has a càdlàg modification which is a Lévy process. As in Schoutens (2003) and Cont and Tankov (2003), the following theory will work with the càdlàg version of the process. Arithmetic Brownian motion and the Poisson process are examples of Lévy processes. It should be noted that the third condition does not imply continuous sample paths; it states that for a given time *t*, the probability of observing a jump at that particular time is zero, so that discontinuities occur at random times.



### Characteristic Functions

The characteristic function $\varphi$ of an $\mathbb{R}^d$-valued random variable $X$ with respect to the probability measure $\mu$ is defined as:

$$\varphi_X(z) = \mathbb{E}[e^{iz\cdot X}] = \int_{\mathbb{R}^d} e^{iz\cdot x} d\mu_x(x)\,, \qquad \forall z \in \mathbb{R}^d. \tag{3.1}$$

The characteristic function is the Fourier transform of the distribution of the random variable. It completely characterizes the law of the random variable *X* so that two random variables with the

same characteristic function have the same distribution. The notation for the characteristic function of the Lévy process $X_t$ is defined accordingly:

$$\varphi_t(z) = \varphi_{X_t}(z) = \mathbb{E}[e^{iz \cdot X_t}], \qquad z \in \mathbb{R}^d. \tag{3.2}$$

Now, for $t > s$, write $X_{t+s} = X_s + (X_{t+s} - X_s)$. Then since $X_{t+s} - X_s$ is independent of $X_s$ and the process has stationary increments:

$$\begin{aligned} \varphi_{t+s}(z) &= \varphi_{X_{t+s}}(z) = \varphi_{X_s}(z)\varphi_{X_{t+s}-X_s}(z) = \varphi_{X_s}(z)\varphi_{X_t}(z) \\ &= \varphi_s(z)\varphi_t(z) \end{aligned}$$

then the stochastic continuity of $t \to X_t$ implies in particular that $X_s \to X_t$ in distribution when $s \to t$. Therefore $\varphi_{X_s}(z) \to \varphi_{X_t}(z)$ when $s \to t$, so that $t \to \varphi_{X_t}(z)$ is a continuous function of $t$. This, together with the multiplicative property $\varphi_{s+t}(z) = \varphi_s(z)\varphi_t(z)$ implies that $t \to \varphi_{X_t}(z)$ is an exponential function and leads us to the following proposition:

**Proposition 3.2 (Characteristic function of a Lévy process).** *(Cont and Tankov, 2003, §3.2) Let $X = (X_t)_{t \geq 0}$ be a Lévy process on $\mathbb{R}^d$. There exists a continuous function $\psi : \mathbb{R}^d \mapsto \mathbb{R}$ called the characteristic exponent of $X$, such that:*

$$\mathbb{E}[e^{iz \cdot X_t}] = e^{t\,\psi(z)}, \qquad z \in \mathbb{R}^d. \tag{3.3}$$

The characteristic exponent (Lévy exponent) completely determines the law of $X$. More specifically, the law of $X$ can determined by specification of the distribution of $X_t$ for some $t$ (for example $t = 1$). Model prices of European vanilla options will be computed via the characteristic function as per section 3.7.3.

## 3.2 Compound Poisson Processes

The Poisson process is a counting process and is an example of a Lévy process. The compound Poisson process will be used in the construction of option pricing models with jumps - the jump terms in the Merton (1976) and Kou (2002) jump diffusion models are defined by compound Poisson processes. The same compound Poisson processes will be use to incorporate jumps into stochastic volatility models, as is done in the Bates (1996) model.

### The Exponential Distribution

The Exponential distribution is closely linked to the Poisson process and is therefore considered first. It is a continuous time distribution which is used to model the time between the occurrences of independent random events which occur with average rate $\lambda$ over a given time period.

**Definition 3.3** (**The Exponential Distribution**). A positive random variable $Y$ is said to follow an exponential distribution with parameter $\lambda > 0$ if it has a probability density function of the form

$$f_Y(y) = \lambda e^{-\lambda y} \mathbf{1}_{y \geq 0}. \tag{3.4}$$

The corresponding distribution function is given by

$$F_Y(y) = \mathbb{P}(Y \leq y) = 1 - e^{-\lambda y}, \qquad \forall y \in [0, \infty] \tag{3.5}$$

with inverse distribution function

$$F_Y^{-1}(y) = -\frac{1}{\lambda} \ln(1 - y), \qquad \forall y \in [0, 1]. \tag{3.6}$$

The inverse distribution function can be used to simulate exponential random variables. If a random number $U \in [0, 1]$ is generated, then $-\frac{1}{\lambda} \ln U$ will follow an exponential distribution with parameter $\lambda$. A well know property of the exponential distribution is the *memoryless* or *absence of memory* property. If T is an exponential random variable, then

$$\begin{aligned} \mathbb{P}(T > t + s | T > t) &= \frac{\mathbb{P}((T > t + s) \cap (T > t))}{\mathbb{P}(T > t)} \\ &= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} \\ &= e^{-\lambda s} \\ &= \mathbb{P}(T > s), \qquad \forall t, s > 0. \end{aligned}$$

Only the Exponential distribution has this property; this leads to the following proposition which is proved in Cont and Tankov (2003):

**Proposition 3.4** (**Absence of memory**). *(Cont and Tankov, 2003, §2.8)*
*Let $T \geq 0$ be a nonzero random variable such that*

$$\mathbb{P}(T > t + s | T > t) = \mathbb{P}(T > s), \qquad \forall t, s > 0. \tag{3.7}$$

*Then $T$ has exponential distribution.*

Under the Kou (2002) model, the jumps in the log-asset price are double exponentially distributed. That is, the sizes of the upward jumps and the sizes of the downward jumps are given by exponential distributions, each with different parameters. Analytic solutions for prices of certain path dependent options (barrier, look-back) are facilitated by the memoryless property of the exponential distribution, as detailed in Kou and Wang (2004).

### The Poisson Distribution

The Poisson distribution is a discrete time distribution which gives the probability of exactly $n$

occurrences of an event over a given time period. $\lambda$ represents the mean number of occurrences of the event over the given time period.

**Definition 3.5** (**The Poisson Distribution**). An integer valued random variable $N$ is said to follow a Poisson distribution with parameter $\lambda$ if

$$\mathbb{P}(N = n) = e^{-\lambda}\frac{\lambda^n}{n!}, \qquad \forall n \in \mathbb{N}. \tag{3.8}$$

The following proposition gives the link between the Poisson distribution and sums of exponentially distributed random variables.

**Proposition 3.6.** *(Cont and Tankov, 2003)*
*If $(\tau_i)_{i\geq 1}$ are independent exponential random variables with parameter $\lambda$ then, for any $t > 0$ the random variable*

$$N_t = \inf\{n \geq 1, \sum_{i=1}^{n} \tau_i > t\} \tag{3.9}$$

*follows a Poisson distribution with parameter $\lambda t$:*

$$\mathbb{P}(N_t = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \qquad \forall n \in \mathbb{N}. \tag{3.10}$$

The characteristic function of the Poisson distribution with parameter $\lambda$ is given by

$$\varphi_N(z) = \mathbb{E}[e^{iz\cdot N}] = e^{\lambda(e^{iz}-1)}, \qquad \forall z \in \mathbb{R}. \tag{3.11}$$

**Poisson Processes**

**Definition 3.7** (**Poisson process**). (Cont and Tankov, 2003, §2.17)
Let $(\tau_i)_{i\geq 1}$ be sequence of independent exponential random variables with parameter $\lambda$ and $T_n = \sum_{i=1}^{n} \tau_i$.
The process $(N_t)_{t\geq 0}$ defined by

$$N_t = \sum_{n\geq 1} \mathbf{1}_{t\geq T_n} \tag{3.12}$$

is called a Poisson process with intensity $\lambda$.

The Poisson process is termed a counting process since it counts the number of random times which occur on $[0, t]$. Note that for $n \geq 1$, $T_n - T_{n-1} = \tau_n$. The Poisson process is used in jump diffusion models to give the number of jumps which occur over a time interval.

**Definition 3.8** (**Compound Poisson process**). (Cont and Tankov, 2003, §3.3)
A compound Poisson process with intensity $\lambda > 0$ and jump size distribution $f$ is a stochastic process $X_t$ defined as

$$X_t = \sum_{i=1}^{N_t} Y_i \quad \left( = \sum_{n\geq 1} Y_n \mathbf{1}_{N_t \geq n} = \sum_{n\geq 1} Y_n \mathbf{1}_{t\geq T_n} \right), \tag{3.13}$$

where jump sizes $Y_i$ are *i.i.d.* with distribution $f$ and $(N_t)$ is a Poisson process with intensity $\lambda$, independent from $(Y_i)_{i\geq 1}$.

The compound Poisson process is a Lévy process. The special case of the compound Poisson process with $Y_i \equiv 1$ gives the Poisson process. The jump processes which are used to define the jump-diffusion models in chapter 4, and the models with stochastic volatility and jumps in chapter 5, are compound Poisson processes. Note that every càdlàg function can be approximated by piecewise constant functions. The following proposition, which is proved in Cont and Tankov (2003), states that the compound Poisson process is the only Lévy process with piecewise constant sample paths. This leads to the idea of approximating Lévy processes using compound Poisson processes.

**Proposition 3.9.** *$(X_t)_{t\geq 0}$ is a compound Poisson process if and only if it is a Lévy process and its sample paths are piecewise constant functions.*

Now, denoting the characteristic function of the jump size distribution $f$ by $\hat{f}(u)$ $(= \mathbb{E}[e^{iu\cdot Y}])$, the characteristic function of the compound Poisson process with jump intensity $\lambda$ is given by

$$\begin{aligned}
\mathbb{E}[e^{iu\cdot X_t}] &= \sum_{n=0}^{\infty} \mathbb{E}[e^{iu\cdot X_t}|N_t = n]\mathbb{P}(N_t = n) \\
&= \sum_{n=0}^{\infty} \mathbb{E}[e^{iu\cdot \sum_{i=1}^{n} Y_i}|N_t = n]\mathbb{P}(N_t = n) \\
&= \sum_{n=0}^{\infty} \left(\mathbb{E}[e^{iu\cdot Y}]\right)^n e^{-\lambda t}\frac{(\lambda t)^n}{n!} \\
&= e^{-\lambda t}\sum_{n=0}^{\infty} \frac{\left(\hat{f}(u)\lambda t\right)^n}{n!} \\
&= e^{-\lambda t}e^{\hat{f}(u)\lambda t} \\
&= e^{\lambda t(\hat{f}(u)-1)}
\end{aligned}$$

this then leads to

**Proposition 3.10 (Characteristic function of a compound Poisson process).** *(Cont and Tankov, 2003, §3.4) Let $(X_t)_{t\geq 0}$ be a compound process on $\mathbb{R}^d$ with $\lambda$ and $f$ defined as above. Then its characteristic function has the following representation:*

$$\mathbb{E}[e^{iu\cdot X_t}] = \exp\left\{t\lambda \int_{\mathbb{R}^d} (e^{iu\cdot x} - 1)f(dx)\right\}, \qquad \forall u \in \mathbb{R}^d. \tag{3.14}$$

If a new measure is then introduced, denoted by $\nu(A) = \lambda f(A)$, then (3.14) can be rewritten as

$$\mathbb{E}[e^{iu\cdot X_t}] = \exp\left\{t \int_{\mathbb{R}^d} (e^{iu\cdot x} - 1)\nu(dx)\right\}, \qquad \forall u \in \mathbb{R}^d \tag{3.15}$$

where $\nu$ is called the Lévy measure of the process (see definition 3.14).

## 3.3 Random Measures and Jump Measures

It is necessary to review some basic theory relating to random measures in this section before encountering fundamental results of Lévy processes in the following section. Recall the Poisson process $(N_t)_{t\geq 0}$, and if $T_1, T_2, \ldots$ is a sequence of jump times of $N$, then $N_t$ is the number of jumps which occur before time $t$:

$$N_t = \#\{i \geq 1, T_i \in [0,t]\} = \sum_{i\geq 1} \mathbf{1}_{T_i \in [0,t]}. \tag{3.16}$$

The $\tau_i$ (or the times between jumps) are independent exponential random variables, so that the jump times themselves (the $T_i$) are random on the interval $[0,\infty]$. The counting procedure of the Poisson process defines a measure $M$ on $[0,\infty]$: for any measurable set $A \subset \mathbb{R}^+$ let

$$M(\omega, A) = \#\{i \geq 1, T_i(\omega) \in A\}. \tag{3.17}$$

First note that the measure $M$ depends on $\omega$ and is therefore termed a random measure, specifically, $M$ is termed the *random jump measure* associated with the Poisson process. From the definition of the Poisson process, $M$ is positive and integer valued. Additionally, $M(A)$ is finite with probability 1 for any bounded set $A$. The average value of the random measure $M$ is determined by the intensity of the Poisson process ($\lambda$):

$$\mathbb{E}[M(A)] = \lambda |A| \tag{3.18}$$

where $|A|$ is the Lebesgue measure of $A$. The Poisson process can be expressed in terms of the random measure $M$:

$$N_t(\omega) = M(\omega, [0,t]) = \int_{[0,t]} M(\omega, ds). \tag{3.19}$$

The random counting measure defined in (3.17) can be extended from domain $\mathbb{R}^+$ to $\mathbb{R}^d$, in which case the Radon measure ($\mu$) will need to be used in place of the Lebesgue measure.

**Definition 3.11 (Poisson random measure).** (Cont and Tankov, 2003, §2.18)
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $E \subset \mathbb{R}^d$ and $\mu$ a given (positive) Radon measure on $(E, \mathcal{E})$. A Poisson random measure on $E$ with intensity measure $\mu$ is a integer valued random measure:

$$\begin{aligned} M : \Omega \times \mathcal{E} &\rightarrow \mathbb{N} \\ (\omega, A) &\mapsto M(\omega, A), \end{aligned} \tag{3.20}$$

such that

(i) For (almost all) $\omega \in \Omega$, $M(\omega, .)$ is an integer valued Radon measure on $E$: for any bounded measurable $A \subset E$, $M(A) < \infty$ is an integer valued random variable.

(ii) For each measurable set $A \subset E$, $M(., A) = M(A)$ is a Poisson random variable with

parameter $\mu(A)$:

$$\mathbb{P}(M(A) = k) = e^{-\mu(A)} \frac{(\mu(A))^k}{k!}, \qquad \forall k \in \mathbb{N}.$$

(iii) For disjoint measurable sets $A_1, \ldots, A_n \in \mathcal{E}$, the variables $M(A_1), \ldots, M(A_n)$ are independent.

Following on from the Poisson random measure, a *jump measure* is now defined for any Lévy process.

**Definition 3.12 (Jump measure).** Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$. Then, a random measure $J : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{N}$ can be associated with $X$: for any measurable set $B \subset \mathbb{R}^d \times \mathbb{R}^+$

$$J_X(B) = \#\{([0,t], \Delta_t) \in B\}, \qquad \Delta_t = X_t - X_{t-} \tag{3.21}$$

$J$ counts the number of jumps of $X$ on the interval $[0,t]$, with jump sizes in $B$. Note that $X$ can be said to have continuous sample paths if and only if $J_X = 0$ almost surely.

The following proposition suggests an interpretation of the Lévy measure of a compound Poisson process as the average number of jumps per unit of time. This idea is used to define the Lévy measure for all Lévy processes in the subsequent definition 3.14.

**Proposition 3.13 (Jump measure of a compound Poisson process).** *(Cont and Tankov, 2003, §3.5) Let $(X_t)_{t\geq 0}$ be a compound Poisson process with intensity $\lambda$ and jump size distribution $f$. Its jump measure $J_X$ is a Poisson random measure on $\mathbb{R}^d \times \mathbb{R}^+$ with intensity measure $\mu(dx \times dt) = \nu(dx)dt = \lambda f(dx)dt$.*

*Proof.* (Cont and Tankov, 2003, §3.5) It follows from the definition in equation 3.21 that $J_X$ is an integer valued measure. Now, first check that $J_X(B)$ is Poisson distributed. It is sufficient to prove this property for a set of the form $B = A \times [t_1, t_2]$ with $A \in \mathcal{B}(\mathbb{R}^d)$. Let $(N_t)_{t\geq 0}$ be the Poisson process, counting the jumps of X. Conditionally on the trajectory of $N$, the jump sizes $Y_i$ are i.i.d. and $J_X([t_1,t_2] \times A)$ is a sum of $N(t_2) - N(t_1)$ i.i.d. Bernoulli variables taking value 1 with probability $f(A)$. Therefore,

$$\begin{aligned}
\mathbb{E}[e^{iuJ_X([t_1,t_2]\times A)}] &= \mathbb{E}[\mathbb{E}[e^{iuJ_X([t_1,t_2]\times A)} | N_t, t \geq 0]] \\
&= \mathbb{E}[\{e^{iu} f(A) + 1 - f(A)\}^{N(t_2)-N(t_1)}] \\
&= e^{\lambda(t_2-t_1)f(A)(e^{iu}-1)}
\end{aligned}$$

since $N(t_2) - N(t_1)$ is Poisson distributed with parameter $\lambda(t_2 - t_1)$. Thus, $J_X([t_1,t_2] \times A)$ is a Poisson random variable with parameter $\lambda f(A)(t_2 - t_1)$.

Next, check the independence of measures of disjoint sets. First, show that if $A$ and $B$ are two disjoint Borel sets in $\mathbb{R}^d$ then $J_X([t_1,t_2] \times A)$ and $J_X([t_1,t_2] \times B)$ are independent. Conditionally on the trajectory of $N$, the expression

$$iuJ_X([t_1,t_2] \times A) + ivJ_X([t_1,t_2] \times B)$$

is a sum of $N(t_2) - N(t_1)$ i.i.d. random variables taking values: $iu$ with probability $f(A)$, $iv$ with probability $f(B)$, 0 with probability $(1 - f(A) - f(B))$. Proceeding as above, the characteristic function is factorized as:

$$\begin{aligned}
&\mathbb{E}[e^{iuJ_X([t_1,t_2]\times A)+ivJ_X([t_1,t_2]\times B)}] \\
&\quad = \mathbb{E}[\{(e^{iu}-1)f(A) + (e^{iv}-1)f(B) + 1\}^{N(t_2)-N(t_1)}] \\
&\quad = \exp\{\lambda(t_2-t_1)(f(A)(e^{iu}-1) + f(B)(e^{iv}-1))\} \\
&\quad = \mathbb{E}[e^{iuJ_X([t_1,t_2]\times A)}]\mathbb{E}[e^{ivJ_X([t_1,t_2]\times B)}]
\end{aligned}$$

Second, let $[t_1, t_2]$ and $[s_1, s_2]$ be two disjoint intervals. The independence of $J_X([t_1, t_2] \times A)$ and $J_X([s_1, s_2] \times B)$ follows directly from the independence of increments of the process $X$. The independence of jump measures of any finite number of disjoint sets of $[0, \infty] \times \mathbb{R}^d$ follows directly from the fact that the methods used in this proof work for any finite number of sets and from the additivity of $J_X$.

□

**Definition 3.14 (Lévy measure).** (Cont and Tankov, 2003, §3.4)
Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$. The measure $\nu$ on $\mathbb{R}^d$ defined by:

$$\nu(A) = \mathbb{E}[\#\{t \in [0,1] : \Delta X_t \neq 0, \Delta X_t \in A\}], \qquad A \in \mathcal{B}(\mathbb{R}^d) \tag{3.22}$$

is called the Lévy measure of $X$ : $\nu(A)$ is the expected number, per unit time, of jumps whose size belongs to $A$.

The definition of the jump measure $J_X$ for the compound Poisson process $X$ leads to the representation of $X$ as

$$X_t = \sum_{s\in[0,t]} \Delta X_s = \int_{[0,t]\times\mathbb{R}^d} x J_X(ds \times dx) \tag{3.23}$$

Note that there are no convergence problems because the stochastic integral is a finite sum; there are a finite number of jumps (a.s.) in the interval $[0, t]$ for any compound Poisson process.

## 3.4 Lévy-Itô Decomposition

The preceding section defined random measures (the Poisson random measure, the jump measure and the Lévy measure); it led to the representation of the compound Poisson process as a stochastic integral for some Poisson random measure. Then, by proposition 3.9, every piecewise constant Lévy process can also be represented as a stochastic integral for some Poisson random measure. Note that the intensity measure of the Poisson random measure takes the form $\nu(dx)dt$, where $\nu$ is a finite (Lévy) measure as in definition 3.14.

Taking a closer look at the Lévy measure $\nu$, note that $\nu(A)$ is finite for a compact subset of $\mathbb{R}^d \backslash \{0\}$; this is because of the imposition of the càdlàg property. That is, if $\nu(A)$ were not finite on this set then there would be a infinite number of finite size jumps, contradicting the càdlàg property.

$\nu$ can still diverge at 0, and therefore is not a finite measure - $X$ may have an infinite number of small jumps, so that the summation of the jumps becomes an infinite series. Therefore, in order to ensure convergence, conditions must be placed on the Lévy measure $\nu$.

The next proposition brings all of these ideas together and gives a decomposition for Lévy processes into four different components (a detailed proof can be found in (Sato, 1999, Chapter 4)).

**Proposition 3.15 (Lévy-Itô decomposition).** *(Cont and Tankov, 2003, §3.7)*
*Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$ and $\nu$ its Lévy measure, given by definition 3.14. If*

- *$\nu$ is a Radon measure on $\mathbb{R}^d\backslash\{0\}$ and verifies:*

$$\int_{|x|\leq 1} |x|^2 \nu(dx) < \infty \qquad \int_{|x|\geq 1} \nu(dx) < \infty. \tag{3.24}$$

- *The jump measure of $X$, denoted by $J_X$, is a Poisson random measure on $[0,\infty]\times\mathbb{R}^d$ with intensity measure $\nu(dx)dt$.*

*Then there exist a vector $\gamma$ and a d-dimensional Brownian motion $(B_t)_{t\geq 0}$ with covariance matrix A such that*

$$X_t = \gamma t + B_t + X_t^l + \lim_{\epsilon\downarrow 0} \tilde{X}_t^\epsilon \tag{3.25}$$

*where*

$$X_t^l = \int_{|x|\geq 1, s\in[0,t]} x J_X(ds\times dx) \tag{3.26}$$

$$\tilde{X}_t^\epsilon = \int_{\epsilon\leq|x|<1, s\in[0,t]} x\{J_X(ds\times dx) - \nu(dx)ds\} \tag{3.27}$$

$$\equiv \int_{\epsilon\leq|x|<1, s\in[0,t]} x\tilde{J}_X(ds\times dx). \tag{3.28}$$

*and the terms in 3.25 are independent, the convergence in the last term is almost sure and uniform in t on $[0,T]$.*

The Lévy-Itô decomposition requires that for every Lévy process there exists a positive definite matrix $A$, a vector $\gamma$ and a positive measure $\nu$ that uniquely determine its distribution. The parameters $(A,\gamma,\nu)$ are called the *characteristic triplet* or *Lévy triplet* of the process $X$.

The first two terms in equation 3.25 represent continuous processes. Their sum, $\gamma t + B_t$, forms a continuous Gaussian Lévy process (an arithmetic Brownian motion with drift). This process is determined only by the parameters $A$ (the covariance matrix of the Brownian motion) and $\gamma$ (the drift term). The third parameter of the triplet, $\nu$ (the Lévy measure), determines the last two terms in equation 3.25; both of these terms represent discontinuous processes and they incorporate the jumps in the process $X$.

The condition $\int_{|x|\geq 1}\nu(dx) < \infty$ means that the process $X$ has a finite number of jumps whose

absolute value is greater than 1. Therefore the sum

$$X_t^l = \sum_{0\leq s\leq t}^{|\Delta X_s|\geq 1} \Delta X_s \tag{3.29}$$

will contain a finite number of terms (a.s.) and $X_t^l$ is a compound Poisson process. There are no convergence issues here. The sum of jumps whose size lies between $\epsilon$ and 1,

$$X_t^\epsilon = \sum_{0\leq s\leq t}^{\epsilon\leq|\Delta X_s|<1} \Delta X_s = \int_{\epsilon\leq|x|<1, s\in[0,t]} x J_X(ds\times dx) \tag{3.30}$$

is a compound Poisson process but with a problem arising due to the fact that $\nu$ can have a singularity at zero; there can be infinitely many small jumps which means that their sum will not necessarily converge. The problem arises as $\epsilon$ tends down to 0. Therefore $X_t^\epsilon$ is replaced by its compensated (centered) version, $\tilde{X}_t^\epsilon$. The centered version subtracts the expected number of jumps at each point in time and is a martingale (shown in Cont and Tankov (2003)) ensuring convergence (a.s.) of the stochastic integral.

Finally stated, the Lévy-Itô decomposition implies that every Lévy process is a combination of a Brownian motion with drift and a possibly infinite sum of compound Poisson processes. A jump-diffusion process is a combination of a Brownian motion with drift and a compound Poisson process, so the proposition implies that any Lévy process can be approximated arbitrarily closely by a jump-diffusion process.



## 3.5 Lévy-Khinchin Representation

As $\epsilon$ tends down to zero, the Lévy-Itô decomposition says that the decomposed process $(\gamma t + B_t + X_t^l + \tilde{X}_t^\epsilon)$ tends to $X_t$ almost surely; this implies convergence in distribution and hence enables the characteristic function of $X_t$ to be represented as:

$$\begin{aligned}\mathbb{E}[e^{iz.X_t}] &= \mathbb{E}[e^{iz.(\gamma t+B_t+X_t^l+\tilde{X}_t^\epsilon)}]\\ &= \mathbb{E}[e^{iz.(\gamma t+B_t)}]E[e^{iz.X_t^l}]E[e^{iz.\tilde{X}_t^\epsilon}] \qquad \text{(by independence of components)},\end{aligned}$$

where the component characteristic functions are

$$\begin{aligned}E[e^{iz.(\gamma t+B_t)}] &= \exp(ti\gamma.z - \frac{t}{2}z.Az)\\ E[e^{iz.X_t^l}] &= \exp\left(t\int_{|x|\geq 1}(e^{iz.x}-1)\nu(dx)\right)\end{aligned}$$

since $X_t^l = \int_{|x|\geq 1, s\in[0,t]} x J_X(ds\times dx)$ is a compound Poisson process. Furthermore, $\tilde{X}_t^\epsilon = \int_{\epsilon\leq|x|<1, s\in[0,t]} x\{J_X(ds\times dx) - \nu(dx)ds\}$ is a compensated compound Poisson processes with

characteristic function given by

$$E[e^{iz.\tilde{X}_t^\epsilon}] = \exp\left(t\int_{\epsilon\leq|x|<1}(e^{iz.x}-1-iz.x)\nu(dx)\right)$$

which converges as $\epsilon \downarrow 0$, giving

$$\mathbb{E}[e^{iz.X_t}] = \exp\left(t\Big(i\gamma.z-\frac{1}{2}z.Az+\int_{\mathbb{R}^d}(e^{iz.x}-1-iz.x\mathbf{1}_{|x|\leq 1})\nu(dx)\Big)\right).$$

So, the Lévy-Khinchin representation below follows as a consequence of the Lévy-Ito decomposition and it gives the representation of the characteristic function of a Lévy process $X_t$ in terms of its characteristic triplet $(A,\nu,\gamma)$. It is is used to define exponential Lévy martingales in proposition 3.22 (ii). Note that (Sato, 1999, Theorem 8.1) establishes the Lévy-Khinchin result prior to proving the Lévy-Ito decomposition.

**Theorem 3.16 (Lévy-Khinchin representation).** *(Cont and Tankov, 2003, §3.1) Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$ with characteristic triplet $(A,\nu,\gamma)$. Then*

$$\mathbb{E}[e^{iz\cdot X_t}] = e^{t\psi(z)}, \quad z\in\mathbb{R}^d \tag{3.31}$$

$$\text{with} \quad \psi(z) = -\frac{1}{2}z.Az+i\gamma.z+\int_{\mathbb{R}^d}\big(e^{iz\cdot x}-1-iz.x\mathbf{1}_{|x|\leq 1}\big)\nu(dx). \tag{3.32}$$

It should be noted that the $\gamma$ parameter is not the drift of the Lévy process $X_t$. In fact, the drift is defined as

$$b=\gamma-\int_{|x|\leq 1}x\nu(dx)$$

and is only defined if $\int_{|x|\leq 1}|x|\nu(dx)<\infty$.

The following propositions can be made based on the characteristic triplet $(A,\nu,\gamma)$:

**Proposition 3.17 (Path characteristics of Lévy processes).** *(Cont and Tankov, 2003, §3.8,§3.9) and (Schoutens, 2003, Ch 5.1.2)*

- *A Lévy process has piecewise constant trajectories if and only if its characteristic triplet satisfies*
$$A=0,\ \int_{\mathbb{R}^d}\nu(dx)<\infty \quad \textit{and} \quad \gamma=\int_{|x|\leq 1}x\nu(dx)$$
- *A Lévy process is of **finite variation** if and only if*
$$A=0 \quad \textit{and} \quad \int_{|x|\leq 1}|x|\nu(dx)<\infty$$
- *A Lévy process is of **finite activity** if and only if*
$$A=0 \quad \textit{and} \quad \int_{|x|\leq 1}\nu(dx)<\infty$$

*that is, there are finitely many jumps on any finite interval.*

- *A Lévy process is of **infinite variation** if and only if*

$$A \neq 0$$

*so that there is a Brownian component which is of infinite variation,*
***or**,*

$$A = 0 \quad \textit{and} \quad \int_{|x|\leq 1} |x|\nu(dx) = \infty$$

*in which case special attention must be paid to small jumps as their sum does not converge (this leads to the need for the compensator term in the Lévy-Itô decomposition).*

with finite variation defined:

**Definition 3.18 (Finite variation).** If $f : [0, \infty) \to \mathbb{R}^d$ is a function, and given the interval $[a, b]$, then the total variation of $f$ over the interval is defined by

$$V(f; [a, b]) = \sup \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|$$

where the supremum is taken over all finite partitions $0 \leq a = t_0 < t_1 < \cdots < t_n = b < \infty$ of the interval. If the total variation is finite, then the function is said to be of finite variation over the interval.

A one-dimensional finite variation function can be decomposed into a difference of two increasing functions. A Lévy process is said to be of finite variation if all its paths are functions of finite variation with probability 1.

The jump-diffusion models of Merton (1976) and Kou (2002) which are evaluated in chapter 7 are all infinite variation models (since they have a Brownian motion driving component). The pure-jump models (Variance-Gamma) are finite variation, infinite activity models.



## 3.6 Subordination

A subordinator is a non-negative, non-decreasing Lévy process (of finite variation).

**Proposition 3.19.** *A 1-dimensional Lévy process $X$ is non-decreasing a.s. ($t > s \Rightarrow X_t > X_s$) if and only if all of the following conditions are satisfied*

- $A = 0$
- $\nu((-\infty, 0]) = 0$
- $\int_0^\infty (x \wedge 1)\nu(dx) < \infty$
- $b \geq 0$

*that is, $X$ has no diffusion component, jumps of only positive magnitude and finite variation, and has positive drift.*

*Proof.* (Cont and Tankov, 2003, §3.10)
(if:) Since the trajectories are non-decreasing, they are of finite variation. Therefore (by proposition 3.17) $A = 0$ and $\int_{|x|\leq 1} |x|\nu(dx) < \infty$. For trajectories to be non-decreasing, there must be no negative jumps, hence $\nu((-\infty, 0]) = 0$. If a function is nondecreasing then after removing some of its jumps, another nondecreasing function is obtained. When all the jumps are removed from a trajectory of $X_t$, a deterministic function $bt$ is obtained which must therefore be nondecreasing. This implies that $b \geq 0$.
(only if:) Under the conditions stated, the process is of finite variation (by proposition 3.17) and therefore equal to the sum of its jumps plus an increasing linear function. For every trajectory, the number of negative jumps on any fixed interval is a Poisson random variable with intensity 0, hence almost surely 0. This means that almost every trajectory is nondecreasing. □

Let $(S_t)_{t\geq 0}$ be a subordinator, then $S_t$ is a positive random variable for all $t$ and can be described (as in Cont and Tankov (2003)) using a Laplace transform instead of a Fourier transform:

If the characteristic triplet of $S$ is given by $(0, \rho, b)$, then the moment generating function of $S_t$ is

$$\mathbb{E}[e^{uS_t}] = e^{t\,l(u)} \quad \forall u \leq 0, \quad \text{where} \quad l(u) = bu + \int_0^\infty (e^{ux} - 1)\rho(dx). \tag{3.33}$$

$l(u)$ is called the Laplace exponent of $S$.

Time is also a non-negative, non-decreasing process. This leads us to the idea of changing time, specifically, the time component of a Lévy process, using a subordinator $S$. This makes time stochastic; and time (in this context) is related to information flow. So that if the stochastic time runs faster than normal time, then the rate of arrival of new information will be greater. The next theorem gives details on the subordination of a Lévy process.

**Theorem 3.20 (Subordination of a Lévy process).** *(Cont and Tankov, 2003, §4.2)*
*On probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$ with characteristic exponent $\Psi(u)$ and triplet $(A, \nu, \gamma)$ and let $(S_t)_{t\geq 0}$ be a subordinator with Laplace exponent $l(u)$ and triplet $(0, \rho, b)$. Define the process $(Y_t)_{t\geq 0}$ for each $\omega \in \Omega$ by $Y(t, \omega) = X(S(t, \omega), \omega)$. Then,*

*(i) $Y$ is a Lévy process and its characteristic function is*

$$\mathbb{E}[e^{iuY_t}] = e^{tl(\Psi(u))}, \tag{3.34}$$

*noting that the characteristic exponent of $Y$ is obtained by composition of the Laplace exponent of $S$ with the characteristic exponent of $X$.*

*(ii) The triplet $(A^Y, \nu^Y, \gamma^Y)$ of $Y$ is given by*

$$A^Y = bA \tag{3.35}$$

$$\nu^Y(B) = b\nu(B) + \int_0^\infty p_s^X(B)\rho(ds), \quad \forall B \in \mathcal{B}(\mathbb{R}^d) \tag{3.36}$$

$$\gamma^Y = b\gamma + \int_0^\infty \rho(ds) \int_{|x|\leq 1} x p_s^X(dx), \tag{3.37}$$

*where $p_t^X$ is the probability distribution of $X_t$.*

*$(Y_t)_{t\geq 0}$ is said to be subordinate to the process $(X_t)_{t\geq 0}$.*

*Proof.* of (i) (Cont and Tankov, 2003, §4.2)
Denote the filtration of $(S_t)_{t\geq 0}$ by $\mathcal{F}_t^S$, with $\mathcal{F}^S \equiv \mathcal{F}_\infty^S$. For every sequence of times $t_0 < t_1 < \ldots < t_n$ the following can be obtained, using the independent increments property of $X$, the Lévy-Khinchin formula for $X$ and the independent increments property of $S$:

$$\begin{aligned}
&\mathbb{E}\left[\prod_{i=1}^n e^{iu_i(X(S_{t_i})-X(S_{t_{i-1}}))}\right] = \mathbb{E}\left\{\mathbb{E}\left[\prod_{i=1}^n e^{iu_i(X(S_{t_i})-X(S_{t_{i-1}}))}\middle|\mathcal{F}^S\right]\right\} \\
&= \mathbb{E}\left\{\prod_{i=1}^n \mathbb{E}\left[e^{iu_i(X(S_{t_i})-X(S_{t_{i-1}}))}\middle|\mathcal{F}^S\right]\right\} = \mathbb{E}\left\{\prod_{i=1}^n e^{(S_{t_i}-S_{t_{i-1}})\Psi(u_i)}\right\} \\
&= \prod_{i=1}^n \mathbb{E}\left\{e^{(S_{t_i}-S_{t_{i-1}})\Psi(u_i)}\right\} = \prod_{i=1}^n \mathbb{E}\left[e^{iu_i(X(S_{t_i})-X(S_{t_{i-1}}))}\right]
\end{aligned}$$

therefore, $Y$ has independent increments. The stationarity of increments can be shown in the same way.

To show that $Y$ is continuous in probability, first observe that every Lévy process is *uniformly* continuous in probability, due to the stationarity of its increments. Further, for every $\varepsilon > 0$ and $\delta > 0$, one can write:

$$P\{|X(S_s) - X(S_t)| > \varepsilon\} \leq P\{|X(S_s) - X(S_t)| > \varepsilon \big| |S_s - S_t| < \delta\} + P\{|S_s - S_t| \geq \delta\}$$

The first term can be made arbitrarily small simultaneously for all values of $s$ and $t$ by changing $\delta$, because $X$ is uniformly continuous in probability. The limit of the second term as $s \to t$ is always zero, because $S$ is continuous in probability. Hence, $P\{|X(S_s) - X(S_t)| > \varepsilon\} \to 0$ as $s \to t$.

Equation 3.34 is obtained by conditioning on $\mathcal{F}^S$:

$$\mathbb{E}[e^{iuY_t}] = \mathbb{E}[e^{iuX(S_t)}] = \mathbb{E}\{\ \mathbb{E}[e^{iuX(S_t)}\ |\mathcal{F}^S]\} = \mathbb{E}[e^{S_t\Psi(u)}] = e^{tl(\Psi(u))}.$$

A detailed proof of (ii) can be found in (Sato, 1999, Theorem 30.1). □

This theorem applies to Brownian motion since it is a Lévy process, and is necessary to develop option pricing models which are based on a subordinated Brownian motion. One such model is the Variance-Gamma pure jump model (section 4.3) which evaluates an arithmetic Brownian motion

at times given by a gamma process; in this case, it should be noted the process formed when the Brownian motion is subordinated is no longer a diffusion process, but a pure jump process. It is possible to go one step further and to model time itself as a stochastic process, so that the time input into the gamma process is stochastic. This is the basis of the VG model with stochastic time (section 4.4).

## 3.7 Exponential Lévy Models

Assume that the dynamics of the asset price are given by

$$S_t = S_0 e^{rt+X_t} \tag{3.38}$$

where $X$ is a Lévy process which accounts for random moves in the asset price. A model with these dynamics is called an exponential Lévy model.

### 3.7.1 Lévy Processes as Martingales

Martingale theory is the basis of risk-neutral pricing, and must be considered in the context of Lévy processes. Key results are taken from Cont and Tankov (2003), beginning with how martingales can be formed from independent increments processes.

**Proposition 3.21.** *Cont and Tankov (2003, §3.17)*
*Let $(X_t)_{t\geq 0}$ be a real-valued process with independent increments, then*

*(i)* $\left(\frac{e^{iuX_t}}{E[e^{iuX_t}]}\right)_{t\geq 0}$ *is a martingale for all $u \in \mathbb{R}$.*

*(ii) If, for some $u \in \mathbb{R}$, $\mathbb{E}[e^{uX_t}] < \infty \forall t \geq 0$ then* $\left(\frac{e^{uX_t}}{E[e^{uX_t}]}\right)_{t\geq 0}$ *is a martingale.*

*(ii) If $\mathbb{E}[X_t] < \infty \forall t \geq 0$ then $M_t = X_t - \mathbb{E}[X_t]$ is a martingale.*

*(iv) If $Var[X_t] < \infty \forall t \geq 0$ then $(M_t)^2 - \mathbb{E}[(M_t)^2]$ is a martingale.*

*If $(X_t)$ is a Lévy process, then for all of the processes (i-iv) given above to be martingale it suffices that the corresponding moments be finite for one value of t.*

This proposition, along with the Lévy-Khinchin formula leads to:

**Proposition 3.22.** *Cont and Tankov (2003, §3.18)*
*Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}$ with characteristic triplet $(A, \nu, \gamma)$,*

*(i) $(X_t)$ is a martingale if and only if $\int_{|x|\geq 1} |x|\nu(dx) < \infty$ and*

$$\gamma + \int_{|x|\geq 1} x\nu(dx) = 0. \tag{3.39}$$

*(ii)* $e^{X_t}$ *is a martingale if and only if* $\int_{|x|\geq 1} e^x \nu(dx) < \infty$ *and*

$$\psi_X(-i) = \gamma + \frac{A}{2} + \int_{\infty}^{\infty} \left(e^x - 1 - x\mathbf{1}_{|x|\leq 1}\right)\nu(dx) = 0. \tag{3.40}$$

Note, following on from (ii) above, if $\psi_X(-i) = 0$ then $\varphi_{X_t}(-i) = 1$. So,

$$\mathbb{E}[e^{X_t}] = \varphi_{X_t}(-i) = e^{t\psi_X(-i)} = 1 \tag{3.41}$$

### 3.7.2 Pricing Vanilla European Options

Recall that the first fundamental theorem of asset pricing (theorem 2.3) states that the pricing model is arbitrage free if an only if there exists an equivalent martingale measure. So, if the risk-neutral measure $\mathbb{Q}$ exists (under which the drift of the asset price process is equal to the risk free rate and the discounted asset price process is a martingale) then the model is arbitrage free. The risk-neutral valuation formula (see theorem 2.4) can then be applied to give the arbitrage free price of an option as the discounted expected value of the payoff under $\mathbb{Q}$:

$$\Pi_t(K,T) = e^{-r(T-t)}\mathbb{E}^{\mathbb{Q}}[g(S_T)|\mathcal{F}_t] \tag{3.42}$$

where $T$ is the time to maturity. $g(S_T) = (S_T - K)^+$ in the case of a call option, where $K$ is the strike price, then the arbitrage free price $C$ at time $t = 0$ is given by

$$C(K,T) = \mathbb{E}^{\mathbb{Q}}[e^{-rT}(S_T - K)^+]. \tag{3.43}$$

**Pricing via the Density Function**

If the density function of the asset price at maturity $T$ under the risk-neutral measure $\mathbb{Q}$ is known, denoted by $f_{\mathbb{Q}}(s,T)$, then the price of a European call option can be calculated by

$$\begin{aligned} C(K,T) &= e^{-rT}\int_0^{\infty} f_{\mathbb{Q}}(s,T)(s-K)^+ ds \\ &= e^{-rT}\int_K^{\infty} f_{\mathbb{Q}}(s,T)(s-K) ds \\ &= e^{-rT}\int_K^{\infty} f_{\mathbb{Q}}(s,T)s\, ds - Ke^{-rT}\mathbb{Q}(s \geq K) \end{aligned} \tag{3.44}$$

The problem with a number of practical models is that the density function is usually not available; this then leads us to work with the characteristic function, which is usually available.

### 3.7.3 Pricing via the Characteristic Function

Recall the asset price dynamics in equation 3.38 which defines $S_t = S_0 e^{rt+X_t}$, where $X$ is a Lévy process. In order to guarantee that the discounted asset price $e^{-rt}S_t$ is a martingale (hence specify

risk-neutral dynamics), it will be necessary to apply proposition 3.22. So, if $X$ has characteristic triplet $(A, \nu, \gamma)$ then we require the following two restrictions to hold:

$$\int_{|x|\geq 1} e^x \nu(dx) < \infty \tag{3.45}$$

$$\psi_X(-i) = \gamma + \frac{A}{2} + \int_{\infty}^{\infty} \big(e^x - 1 - x\mathbf{1}_{|x|\leq 1}\big)\nu(dx) = 0. \tag{3.46}$$

then $\mathbb{E}[e^{-rt}S_t] = \mathbb{E}[S_0 e^{-rt} e^{rt+X_t}] = S_0 \mathbb{E}[e^{X_t}] = S_0 e^{t\psi_X(-i)} = S_0$.

Then the value of a vanilla European option can be computed using Fourier inversion. The following formulae give European call option prices through integrals in Fourier space; this is termed the direct integration method. The Heston (1993) model was the first stochastic volatility model to use this method; the resulting formulae resembled the form of the Black-Scholes formulae. Bakshi and Madan (2000) then generalized this approach. Lewis (2001) then developed an alternative formula which required only one numerical integration instead of two. The Lewis (2001) formula is used in Gatheral (2006) and will be the version used in this dissertation:

**Direct Integration Formula - version 1** (Lewis, 2001)

$$C(S,K,T) = Se^{-qT} - \frac{1}{\pi}\sqrt{SK}e^{-(r+q)T/2}\int_0^{\infty} \frac{\mathbf{Re}\Big[e^{iuk}\varphi_T(u-\frac{i}{2})\Big]}{u^2+\frac{1}{4}}\,du \tag{3.47}$$

where

$$k = \log\Big(\frac{S}{K}\Big) + (r-q)T, \qquad S = S_0$$

**Direct Integration Formula - version 2**

This alternative version is presented in Attari (2004) and is used in Kilin (2007).

$$C(S,K,T) = S - \frac{e^{-rT}K}{2} -$$
$$e^{-rT}K\frac{1}{\pi}\int_0^{\infty} \frac{\big(\mathbf{Re}(\varphi(u)) + \frac{\mathbf{Im}(\varphi(u))}{u}\big)cos(ul(K)) + \big(\mathbf{Im}(\varphi(u)) - \frac{\mathbf{Re}(\varphi(u))}{u}\big)sin(ul(K))}{1+u^2}\,du$$

where

$$l(K) = \ln\Big(\frac{Ke^{-(r-q)T}}{S}\Big), \qquad S = S_0$$

**Fast Fourier Transform**

Another method that allows pricing through the characteristic function was presented in Carr and Madan (1998) and used in Schoutens (2003); it requires a dampening parameter $\alpha$ to be specified. Then

$$C(K,T) = \frac{e^{-\alpha\log(K)}}{\pi}\int_0^{\infty} e^{-iv\log(K)}\varrho(v)dv \tag{3.48}$$

where

$$\varrho(v) = \frac{e^{-rT}\varphi(v-(\alpha+1)i)}{\alpha^2+\alpha-v^2+i(2\alpha+1)v}$$

and the fast Fourier transform (FFT) is used for inversion. This method actually works with a modified call price, where $\alpha$ is the modifying parameter. This final method has been shown to be faster than naive direct integration method due to the fact that one run of the algorithm calculates option prices for a range of strike prices for a specified maturity. Implementation of the FFT method not only requires selection of the $\alpha$ parameter but the method also works with a grid across strikes, therefore, it must be ensured that this grid corresponds to the strikes prices for which option prices are needed. The initial overhead in implementation may only be justified if call option prices are required for a large number of strikes (perhaps over 10), as noted by Cont and Tankov (2003). Kilin (2007) argues that optimized implementation of the direct integration method can be as competitive as the FFT method, with respect to computational time.

# Chapter 4

# Jump Models

Two main types of exponential Lévy models are considered. The first type contains both a diffusion component as well as a jump component and are termed *jump-diffusion models*. The form of the driving Lévy process can be given as

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i \tag{4.1}$$

where $(N_t)_{t\geq 0}$ is a Poisson counting process giving the number of jumps of $X$ and the $Y_i$ are i.i.d. random variables giving the jump sizes. The second type contains a driving Lévy process, $X_t$, which is of infinite jump activity. These models are termed *pure jump models* and are discontinuous everywhere, with no diffusion component so that the process only changes when a jump occurs.



## 4.1 The Kou (2002) Jump-Diffusion Model

The Kou (2002) model is a double exponential jump diffusion model; the model produces a return distribution which is leptokurtic. Empirical observations (as in Cont (2001)) show that asset return distributions are leptokurtic and it is therefore more appropriate for an option pricing model to take this into account. The Model assumes the following real world (P measure) dynamics for the security price $S_t$:

$$\frac{dS_t}{S_{t-}} = \mu dt + \sigma dW_t + d\left(\sum_{i=1}^{N_t}(V_i - 1)\right) \tag{4.2}$$

with solution:

$$S_t = S_0 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W_t + \sum_{i=1}^{N_t} \log V_i\right\} \tag{4.3}$$

where

- $W_t$ is a standard Brownian Motion,
- $N_t$ is a Poisson Process with rate $\lambda$ and

- $\{V_i\}$ is a sequence of independent identically distributed (i.i.d.) random variables with distribution given by 4.4.
- All sources of randomness $(W_t, N_t, Y)$ are assumed to be independent.

$Y = \log(V)$ is the distribution of jump sizes and has an asymmetric double exponential distribution with density

$$f_Y(y) = p.\eta_1 e^{-\eta_1 y}\mathbf{1}_{\{y\geq 0\}} + q.\eta_2 e^{\eta_2 y}\mathbf{1}_{\{y<0\}} \tag{4.4}$$
$$\eta_1 > 1, \eta_2 > 0, p \geq 0, q \geq 0, p+q = 1$$

$p$ and $q$ represent the probabilities of upward and downward jumps, respectively. So,

$$\log(V) = Y =^d \left\{ \begin{array}{ll} \xi^+ & \text{with probability } p \\ -\xi^- & \text{with probability } q \end{array} \right\} \tag{4.5}$$

where $\xi^+ \sim \exp(\frac{1}{\eta_1})$ (i.e. $\xi^+$ follows an exponential distribution with mean $\frac{1}{\eta 1}$) and $\xi^- \sim \exp(\frac{1}{\eta_2})$. Also note that

$$\begin{aligned} \mathbb{E}(V) &= \mathbb{E}(e^Y) \\ &= q\frac{\eta_2}{\eta_2+1} + p\frac{\eta_1}{\eta_1-1}, \qquad \eta_1 > 1, \eta_2 > 0 \end{aligned} \tag{4.6}$$

**Characteristic Function**

The form of the characteristic function for the Kou (2002) model is given in Cont and Tankov (2003) (and Kou and Wang (2004)) as

$$\varphi_T(u) = e^{\psi(u).T} = \exp\left[i\omega uT - \frac{1}{2}\sigma^2 u^2 T + iu\lambda T\Big(\frac{p}{\eta_1 - iu} - \frac{1-p}{\eta_2 + iu}\Big)\right] \tag{4.7}$$

where $\psi$ is the characteristic exponent. To find $\omega$, impose $\varphi_T(-i) = 1$ (to ensure the discounted asset price is a martingale, as per section 3.7.3) giving

$$\begin{aligned} &\exp\left[\omega T + \frac{1}{2}\sigma^2 T + \lambda T\Big(\frac{p}{\eta_1 - 1} - \frac{1-p}{\eta_2+1}\Big)\right] = 1 \\ \Rightarrow \qquad &\omega = -\frac{1}{2}\sigma^2 - \lambda\Big(\frac{p}{\eta_1 - 1} - \frac{1-p}{\eta_2+1}\Big). \end{aligned} \tag{4.8}$$

## Simulation

**Discretization**

The solution of the asset price SDE gives asset price dynamics

$$S_t = S_0 e^{\mu^{\mathbb{Q}} t + \sigma W_t + \sum_{j=1}^{N_t} Y_j}, \tag{4.9}$$

where the drift term $\mu^{\mathbb{Q}}$ ensures that the process is martingale and is defined as

$$\mu^{\mathbb{Q}} = r - q - \frac{1}{2}\sigma^2 - \lambda\Big[\frac{p}{\eta_1 - 1} - \frac{1-p}{\eta_2 + 1}\Big]. \tag{4.10}$$

The logarithm of the security price process is simulated and the discretization scheme, assuming a constant time increment $\Delta t$, is

$$\log S_{i+1} \quad = \quad \log S_i + \mu^{\mathbb{Q}}\Delta t + \sigma\sqrt{\Delta t}.Z + \sum_{j=1}^{N_{\Delta t}} Y_j \tag{4.11}$$

Simulating the Kou jump-diffusion model over time interval $\Delta t$ therefore requires the following steps:

- Generate a standard normal random variable $Z$ to simulate the Brownian motion (diffusion component) on the interval.
- Generate a Poisson random variable $N_{\Delta t}$ with parameter $\lambda\Delta t$ to give the number of jumps on the interval,
- Generate the jumps $Y_j$ from the double exponential distribution which is defined in (4.4). Then the summation of the jumps simulates the jump component on the interval.

**Sampling from the Poisson Distribution**

The cumulative distribution function, F, of the Poisson process is constant on each half open interval $[N, N+1)$. F has an inverse since it is non-decreasing, however, the inverse is not unique on any half open interval. Therefore, F is chosen such that $F^{-1}(u) = \inf\{t|F(t) \geq u\}$, with the appropriate algorithm.

**Sampling from the Asymmetric Double Exponential Distribution**

It is necessary to sample $Y = \log V$ from the distribution defined in (4.4). Firstly,

$$\begin{aligned} &\text{If} && F(x) = 1 - e^{-\eta x} \\ &\text{Then} && F^{-1}(x) = -\frac{1}{\eta}\log(1-U) \\ &\text{Where} && U \sim \text{Uniform}(0,1) \end{aligned} \tag{4.12}$$

Now, in order to sample from a double exponential distribution, it is necessary to draw from

$$\begin{aligned} \xi^+ &\sim \exp(\frac{1}{\eta_1}) && \text{with probability } p \\ -\xi^- &\sim \exp(\frac{1}{\eta_2}) && \text{with probability } q \end{aligned}$$

If $\eta_1$ and $\eta_2$ are substituted for $\eta$ in (4.12) then it is possible to sample from the required exponential distribution. The required distribution is decided upon by drawing another random number $U_2$ and sampling from

$$\begin{aligned} \xi^+ \quad & \text{if} \quad 0 < U_2 < p \\ -\xi^- \quad & \text{if} \quad p < U_2 < 1 \end{aligned}$$

## 4.2 The Merton (1976) Jump-Diffusion Model

The Merton (1976) jump diffusion model assumes that jumps in the log-price are normally distributed with mean jump size $\alpha$ and standard deviation of jump size $\delta$. This gives dynamics

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t + \sum_{i=1}^{N_t} Y_i} \tag{4.13}$$

where

$$Y_i \sim N(\alpha, \delta^2)$$

The Lévy density is given by

$$\nu_X(x) = \frac{\lambda}{\sqrt{2\pi\delta^2}} \exp\left\{ -\frac{(x-\alpha)^2}{2\delta^2} \right\} \tag{4.14}$$

and the characteristic function by

$$\varphi_T(u) = \exp\left\{ iu\omega T - \frac{1}{2}u^2\sigma^2 T + \lambda T\big(e^{iu\alpha - u^2\delta^2/2} - 1\big) \right\}. \tag{4.15}$$

$\omega$ is determined by imposing $\varphi_T(-i) = 1$ (to ensure the discounted asset price is a martingale, as per section 3.7.3) giving

$$\omega = -\frac{1}{2}\sigma^2 - \lambda\Big(e^{\alpha + \delta^2/2} - 1\Big) \tag{4.16}$$

European call options can then be priced via the characteristic function.

**Simulation**

Simulation of the Merton model is similar to that of the Kou model. The solution of the asset price SDE gives asset price dynamics

$$S_t = S_0 e^{\mu^{\mathbb{Q}} t + \sigma W_t + \sum_{j=1}^{N_t} Y_j}, \tag{4.17}$$

where the drift term $\mu^{\mathbb{Q}}$ ensures that the process is martingale and is now defined as

$$\mu^{\mathbb{Q}} = r - q - \frac{1}{2}\sigma^2 - \lambda\big[e^{\alpha - \frac{\delta^2}{2}} - 1\big]. \tag{4.18}$$

The logarithm of the security price process is discretized as

$$\log S_{i+1} \quad = \quad \log S_i + \mu^{\mathbb{Q}}\Delta t + \sigma\sqrt{\Delta t}.Z + \sum_{j=1}^{N_{\Delta t}} Y_j \tag{4.19}$$

This time the $Y_j$'s are normally distributed, thus requiring the following simulating steps for time interval of length $\Delta t$:

- Generate a standard normal random variable $Z$ to simulate the Brownian motion (diffusion component) on the interval.
- Generate a Poisson random variable $N_{\Delta t}$ with parameter $\lambda.\Delta t$ to give the number of jumps on the interval,
- Generate the jumps $Y_j$ from a Normal$(\alpha,\ \delta^2)$ distribution. Then the summation of the jumps simulates the jump component on the interval.

## 4.3 The Variance-Gamma Pure Jump Model

The Variance-Gamma (VG) option pricing model (Madan et al. (1998)) is an exponential Lévy model (see section 3.7.3) which uses a Variance-Gamma process to model random fluctuations in the asset price. The VG process is obtained by evaluating a Brownian motion with drift at random times given by a gamma process. It is a pure jump (no diffusion component) infinite activity process, so that there are an infinite number of jumps in any time period. Intuitively, this ensures that the process better represents actual asset price path behavior which is arguably purely jump driven, with no diffusion component. Also, the process is of finite variation. This can be contrasted to the Brownian motion process which is of finite activity, with infinite variation (but finite quadratic variation).

Define a Brownian motion with drift $\theta$ and volatility $\sigma$ as

$$b(t;\theta,\sigma) = \theta t + \sigma W(t) \tag{4.20}$$

where $W(t)$ is a standard Brownian motion. The gamma process is used as a subordinator.

**Definition 4.1** (**The Gamma process**). If $f$ is a Gamma distribution with parameters $(a, b)$, then the density function is given by

$$f(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-xb}, \qquad a > 0, b > 0, x > 0$$

with characteristic function

$$\varphi(u) = (1 - \frac{iu}{b})^{-a}.$$

Then the Gamma process $(X_t^\gamma)_t \geq 0$ with parameters $(a, b)$ is a stochastic process with stationary and independent Gamma distributed increments. $X_t^\gamma \sim$ Gamma$(at, b)$ and

$$\varphi_t(u) = (1 - \frac{iu}{b})^{-at}.$$

The Gamma process has Lévy density given by

$$\nu_X(x) = \frac{ae^{-bx}}{x} \mathbf{1}_{x>0}.$$

The Gamma process has an infinite arrival rate of jumps (since the Lévy measure has an infinite integral), most of which are small. The Gamma distribution has the scaling property which states that if $f \sim$ Gamma $(a, b)$ then $cf \sim$ Gamma$(a, \frac{b}{c})$.

Now, the VG process is obtained by evaluating a Brownian motion with drift at times given by a Gamma process. So the VG process $X^{VG}(t; \sigma, \nu, \theta)$ is defined in terms of the Brownian motion $b(t; \theta, \sigma)$ and the Gamma process $X^\gamma(t; 1, \nu)$ with unit mean rate (parameters $a = 1/\nu$, $b = 1/\nu$) as

$$\begin{aligned} X^{VG}(t; \sigma, \nu, \theta) &= b(\, X^\gamma(t; 1, \nu);\, \theta,\, \sigma\,) \\ &= \theta X^\gamma(t; \nu) + \sigma W[X^\gamma(t; \nu)]. \end{aligned} \tag{4.21}$$

The VG process has Lévy measure given by

$$\nu_x(x)dx = \frac{\exp(\frac{\theta x}{\sigma^2})}{\nu|x|} \exp\left( - \frac{\sqrt{\frac{2}{\nu} + \frac{\theta^2}{\sigma^2}}}{\sigma} |x| \right) dx \tag{4.22}$$

The VG process can also be expressed as the difference of two independent Gamma processes:

$$X^{VG}(t; \sigma, \nu, \theta) =^d X^{\gamma p}(t; \mu_p, \nu_p) - X^{\gamma n}(t; \mu_n, \nu_n) \tag{4.23}$$

where

$$\mu_p = \frac{\eta_p}{\nu} \qquad \mu_n = \frac{\eta_n}{\nu} \tag{4.24}$$

$$\nu_p = \mu_p^2 \nu \qquad \nu_n = \mu_n^2 \nu \tag{4.25}$$

$$\eta_p = \sqrt{\frac{\theta^2\nu^2}{4} - \frac{\sigma^2\nu}{2}} + \frac{\theta\nu}{2} \tag{4.26}$$

$$\eta_n = \sqrt{\frac{\theta^2\nu^2}{4} + \frac{\sigma^2\nu}{2}} - \frac{\theta\nu}{2} \tag{4.27}$$

allowing the Lévy measure for $X^{VG}$ to be specified as

$$\nu_X(x)dx = \begin{cases} \frac{\mu_n^2}{\nu_n} \frac{\exp\left(-\frac{\mu_n}{\nu_n}|x|\right)}{|x|} dx & \text{for } x < 0 \\ \frac{\mu_p^2}{\nu_p} \frac{\exp\left(-\frac{\mu_p}{\nu_p}|x|\right)}{|x|} dx & \text{for } x > 0 \end{cases}$$

As noted in Schoutens (2003), the Lévy measure has infinite mass ($\int \nu_X(dx) = \infty$) so that a VG process has infinitely many jumps in any finite interval. Also, a VG process has paths of finite variation since $\int |x|\nu_X(dx) < \infty$.

**Characteristic Function**

The characteristic function in the VG model is given by

$$\varphi_T(u) = \exp(iu\omega T)(1 - iu\nu\theta + \frac{1}{2}\sigma^2\nu u^2)^{-\frac{T}{\nu}} \tag{4.28}$$

To find $\omega$, impose $\varphi_T(-i) = 1$ (to ensure the discounted asset price is a martingale, as per section 3.7.3) giving

$$\begin{aligned} e^{\omega T}(1 - \theta\nu - \frac{1}{2}\sigma^2\nu)^{-\frac{T}{\nu}} &= 1 \\ -\frac{T}{\nu}\log(1 - \theta\nu - \frac{1}{2}\sigma^2\nu) &= -\omega T \\ \Rightarrow \omega &= \frac{1}{\nu}\log(1 - \theta\nu - \frac{1}{2}\sigma^2\nu). \end{aligned} \tag{4.29}$$

**Simulation**

The VG model has risk neutral dynamics given by

$$S_t = S_0 e^{\mu^{\mathbb{Q}} t + X_t} \tag{4.30}$$

where

$$\mu^{\mathbb{Q}} = r - q + \frac{1}{\nu}\log\left(1 - \theta\nu - \frac{\sigma^2\nu}{2}\right) \tag{4.31}$$

and $X_t$ is a VG process which can be simulated as the difference of two independent gamma processes, or as a subordinated Brownian motion. The latter approach is used, giving dynamics

$$S_t = S_0 e^{\mu^{\mathbb{Q}} t + \theta G_t + \sigma W_{G_t}}, \tag{4.32}$$

and discretization

$$\log S_{i+1} = \log S_i + \mu^{\mathbb{Q}}\Delta t + \theta G + \sigma\sqrt{G}.Z \tag{4.33}$$

Requiring:

- Generate a standard normal random variable $Z$ to simulate the Brownian motion.
- Generate a Gamma $(\frac{\Delta t}{\nu}, \frac{1}{\nu})$ random variable to simulate the time of the Brownian motion.

### 4.3.1 Extension to the CGMY model

The CGMY is an extension of the VG model, with one extra parameter, $Y$. The characteristic function is

$$\varphi_t^{CGMY}(u) = \exp(i\omega ut)\exp\left\{C\,t\Gamma(-Y)\left[(M-iu)^Y - M^Y + (G+iu)^Y - G^Y\right]\right\} \quad (4.34)$$

where $\omega$ is given by

$$\omega = -C\Gamma(-Y)\left[(M-1)^Y - M^Y + (G+1)^Y - G^Y\right]. \quad (4.35)$$

The Lévy measure for the process is given by

$$\nu_X(x)dx = \begin{cases} C\exp(Gx)(-x)^{-1-Y}dx & \text{for } x < 0 \\ C\exp(-Mx)x^{-1-Y}dx & \text{for } x > 0 \end{cases}$$

It is necessary that the parameters $C\,, G\,, M$ are restricted to $(0, \infty)$ and the following properties follow (from Carr et al. (2002)) for the restricted range of $Y$:

| Range of $Y$ | Process Properties |
|---|---|
| $Y < -1$ | Not completely monotone, finite activity |
| $-1 < Y < 0$ | Completely monotone, finite activity |
| $0 < Y < 1$ | Completely monotone, infinite activity, finite variation |
| $1 < Y < 2$ | Completely monotone, infinite variation, finite quadratic variation |

## 4.4 Variance-Gamma with Stochastic Time

Modeling time as a stochastic process is a method that has been used to incorporate stochastic volatility into Lévy models, see Carr et al. (2001), Schoutens (2003), Cont and Tankov (2003). The idea is to capture random changes in volatility by random changes in time; the Brownian scaling property relates changes in scale to changes in time and is where the idea of this approach stems from. This section will incorporate stochastic time (and hence stochastic volatility effects) into the VG model. There are two main choices of process to use as a stochastic clock: the Cox-Ingersoll-Ross (CIR) process, or the Gamma-Ornstein-Uhlenbeck (G-OU) process. CIR time will be incorporated into the VG model. The approach of Carr et al. (2001) is followed.

**VG(C,G,M)**

The CGM parametrization of the VG model will be worked with, which is equivalent to the CGMY model without the Y parameter. Correspondence between the $VG(\sigma, \nu, \theta)$ and $VG(C, G, M)$ is given by equations 4.26 - 4.27 and

$$C = \frac{1}{\nu} \qquad G = \frac{1}{\eta_n} \qquad M = \frac{1}{\eta_p}. \tag{4.36}$$

The process has characteristic function given by

$$\varphi_{VG}(u; C, G, M) = \left( \frac{GM}{GM + (M - G)iu + u^2} \right)^C \tag{4.37}$$

with corresponding log characteristic function given by

$$\psi_{VG}(u; C, G, M) = C \log \left( \frac{GM}{GM + (M - G)iu + u^2} \right). \tag{4.38}$$

**CIR Time**

Recall that the Heston volatility process was modeled using a mean-revering CIR process, the same one that will be used here to model time. The process $y(t)$ is defined as the solution to the SDE:

$$dy = \kappa(\eta - y)dt + \lambda\sqrt{y}dW \tag{4.39}$$

where

- $W$ is a standard Brownian motion,
- $\eta$ is the long-run rate of time change,
- $\kappa$ is the rate of mean reversion,
- $\lambda$ is the volatility of changes in time.

The new clock is given by the integral of the rate of change of time ($y(t)$):

$$Y(t) = \int_0^t y(u)du. \tag{4.40}$$

This section will use the form of the characteristic function, as given in Cont and Tankov (2003) and Carr et al. (2002):

$$\begin{aligned} \varphi_{CIR}(u, t, y(0); \kappa, \eta, \lambda) &= \mathbb{E}[e^{iuY(t)}] \\ &= A(t, u)e^{B(t,u)y(0)} \end{aligned} \tag{4.41}$$

where

$$A(t,u) = \frac{\exp\left(\frac{\kappa^2\eta t}{\lambda^2}\right)}{\left(\cosh(\frac{\gamma t}{2}) + \frac{\kappa}{\gamma}\sinh(\frac{\gamma t}{2})\right)^{\frac{2\kappa\eta}{\lambda^2}}} \tag{4.42}$$

$$B(t,u) = \frac{2iu}{\kappa + \gamma\coth(\frac{\gamma t}{2})} \tag{4.43}$$

$$\gamma = \sqrt{\kappa^2 - 2\lambda^2 iu}. \tag{4.44}$$

**The VG Stochastic Volatility Process**

The class of stochastic volatility Lévy processes is defined by

$$Z(t) = X(Y(t)), \tag{4.45}$$

where $Y$ is independent of $X$. $Z$ is obtained by subordinating $X$ to $Y$. The characteristic function of process $X$ is given by

$$\varphi_X(u) = e^{t\psi_X(u)} = \mathbb{E}[e^{iuX(t)}]. \tag{4.46}$$

Here $X$ is the VG process and (4.46) is given by (4.37). Recall that $\varphi_{CIR}(u,t,y(0);\kappa,\eta,\lambda) = \mathbb{E}[e^{iuY(t)}]$, then the characteristic function of the process $Z$ is given by

$$\begin{aligned}\varphi_Z(u) &= \mathbb{E}[e^{iuZ(t)}] = \mathbb{E}[e^{Y(t)\psi_X(u)}] \\ &= \varphi_{CIR}\big(-i\psi_X(u), t, y(0);\kappa,\eta,\lambda\big).\end{aligned} \tag{4.47}$$

The VG process with stochastic volatility (VG-CIR) is therefore defined as

$$Z_{VG}(t) = X_{VG}\big(Y(t);C,G,M\big) \tag{4.48}$$

and the parameter $C$ is identified with $y(0)$ in Carr et al. (2001) to give

$$\varphi_Z(u) = \mathbb{E}[e^{iuZ_{VG}(t)}] = \varphi_{CIR}\big(-i\psi_{VG}(u;1,G,M), t, C;\kappa,\eta,\lambda\big), \tag{4.49}$$

leading to a six parameter process.

**Characteristic Function**

The characteristic function of the VG-CIR process can be used to price vanilla European options. The risk-neutral asset price process is given by mean correcting the exponential of the stochastic volatility Lévy process. The asset price at time t is defined by the random variable:

$$S(t) = S(0)\frac{e^{(r-q)t+Z(t)}}{\mathbb{E}[e^{Z(t)}]} \tag{4.50}$$

where $\mathbb{E}[e^{Z(t)}] = \varphi_{CIR}\big( \ -i\psi_{VG}(-i) \ , t, y(0); \kappa, \eta, \lambda\big)$. Note that the factor $\frac{e^{(r-q)t}}{\mathbb{E}[e^{Z(t)}]}$ moves the dynamics to the risk-neutral world by a mean-correcting argument.

The mean corrected characteristic function for the log of the asset price is given by

$$\mathbb{E}[e^{iu\log(S(t))}] = e^{iu\log(S(0))+iu(r-q)t} \cdot \frac{\varphi_{CIR}\big( \ -i\psi_{VG}(u) \ , t, y(0); \kappa, \eta, \lambda\big)}{\varphi_{CIR}\big( \ -i\psi_{VG}(-i) \ , t, y(0); \kappa, \eta, \lambda\big)^{iu}} \tag{4.51}$$

# Chapter 5

# Stochastic Volatility Models with Jumps

A variety of models have been considered so far. These included diffusion models with volatility as a stochastic process, jump-diffusion models with constant diffusion volatility but inclusion of a jump process, pure jump models (which were formed via Brownian subordination) and pure jump models with stochastic time . The final class of models to be reviewed will be of a jump-diffusion type, volatility will follow a stochastic process and jumps in the asset price will be included; termed stochastic volatility with jumps in the underlying (SVJ) models. The diffusion component is assumed to be independent from the jump component, this will allow the characteristic function to be decomposed as

$$\varphi_{X_t}(u) = \varphi_{X_t^c}(u)\varphi_{X_t^j}(u) \tag{5.1}$$

where $\varphi_{X_t^c}$ is the characteristic function of the continuous component, and $\varphi_{X_t^j}$ is the characteristic function of the jump component. The two models considered here will assume the same dynamics for the diffusion component; those being the dynamics assumed in the Heston diffusion model. The version of the characteristic function used is given by Albrecher et al. (2006):

$$\begin{aligned}\varphi_{X_T^c} &= \exp\left\{\frac{\kappa\theta}{\sigma^2}\left\{(\kappa-\rho\sigma iu-d)T-2\log\left[\frac{1-ge^{-dT}}{1-g}\right]\right\}\right\} \\ &\quad\times\exp\left\{\frac{v}{\sigma^2}\big(\kappa-\rho\sigma iu-d\big)\left[\frac{1-e^{-dT}}{1-ge^{-dT}}\right]\right\}\end{aligned} \tag{5.2}$$

where

$$d = \sqrt{\big(\rho\sigma iu-\kappa\big)^2+\sigma^2\big(iu+u^2\big)} \tag{5.3}$$

$$g = \frac{\kappa-\rho\sigma iu-d}{\kappa-\rho\sigma iu+d} \tag{5.4}$$

with

| | |
|---|---|
| $\kappa$ | speed of mean reversion |
| $\theta$ | long variance |
| $\rho$ | correlation |
| $\sigma_v$ | volatility of volatility |
| $v_0$ | short variance |

## 5.1 The Bates (1996) Jump-Diffusion Model

The Bates (1996) model is essentially a Heston-type model with the addition of log-normal jumps in the underlying (Merton-type jumps). The process dynamics are given by

$$\frac{dS_t}{S_t} = (r - \lambda\bar{k})dt + \sqrt{v_t}dW_t^S + dZ_t \tag{5.5}$$

$$dv_t = \kappa(\theta - v_t) + \sigma_v\sqrt{v_t}dW_t^v \tag{5.6}$$

where

- $W_t^S$ and $W_t^v$ are Brownian motions with correlation $\rho$.
- $Z_t$ is a compound Poisson process with intensity $\lambda$ and log-normal distribution of jump sizes,
- $k$ denotes the size of a jump where $\ln(1+k) \sim N(\ln(1+\bar{k}) - \frac{1}{2}\delta^2\,,\,\delta^2)$.
- Jumps are independent of $W^S$ and $W^v$.

The equation for $\log(S_t)$ is given by Itô's formula as

$$dX_t = (r - \lambda\bar{k} - \frac{1}{2}v_t)dt + \sqrt{v_t}dW_t^S + d\tilde{Z}_t \tag{5.7}$$

where $d\tilde{Z}_t$ is a compound Poisson process with intensity $\lambda$ and jump sizes following a Normal distribution. The model parameters can be grouped according to whether the are related to the diffusion (continuous) or the jump component:

| Diffusion | | Jump | |
|---|---|---|---|
| $\kappa$ | speed of mean reversion | $\lambda$ | poisson intensity |
| $\theta$ | long variance | $\bar{k}$ | log-normal jump parameter |
| $\rho$ | correlation | $\delta$ | log-normal jump variance |
| $\sigma_v$ | volatility of volatility | | |
| $v_0$ | short variance | | |

**Characteristic Function**

The specification of the characteristic function is given by equation 5.1 where $\varphi_{X_t^c}$ (equation 5.2) is the characteristic function of the continuous component and $\varphi_{X_t^j}$ is the characteristic func-

tion of the jump component:

$$\varphi_{X_T^j} = \exp\left\{iu\omega T + \lambda T\left(e^{iu\alpha - \frac{u^2\delta^2}{2}} - 1\right)\right\} \tag{5.8}$$

$$\omega = -\lambda\left(e^{\alpha + \frac{\delta^2}{2}} - 1\right) \tag{5.9}$$

$$\alpha = \ln(1+\bar{k}) - \frac{\delta^2}{2} \tag{5.10}$$

$$\Rightarrow \varphi_{X_t^j} = \exp\left\{-iuT\lambda\bar{k} + \lambda T\left(e^{iu(\ln(1+\bar{k}) - \frac{\delta^2}{2}) - \frac{u^2\delta^2}{2}} - 1\right)\right\}. \tag{5.11}$$

**Simulation**

Simulating an SVJ model requires simulation of the volatility process (which is a diffusion process), simulation of the diffusion component of the asset price process and simulation of the jump component of the asset price process. The volatility process is assumed to follow the same mean-reverting process as in the Heston model, recall the discretization from 2.29:

$$v_{i+1} = \left(\sqrt{v_i} + \frac{\sigma_v}{2}\sqrt{\Delta t}Z^v\right)^2 + \kappa(\theta - v_i)\Delta t - \frac{\sigma_v^2}{4}\Delta t \tag{5.12}$$

and define the risk-neutral drift term by

$$\mu_i^{\mathbb{Q}} = r - q - \lambda\bar{k} - \frac{1}{2}v_i, \tag{5.13}$$

then the logarithm of the asset price can be discretized as

$$\log S_{i+1} = \log S_i + \mu_i^{\mathbb{Q}}\Delta t + \sqrt{v_i\Delta t}\; Z^S + \sum_{j=1}^{N_{\Delta t}} Y_j \tag{5.14}$$

Therefore, simulating the asset price over time step $\Delta t$ would involve the following steps:

- Generate two standard normal random variables $Z^S$ and $Z$; where $Z^S$ generates the Brownian motion which drives the asset price process, $Z$ is independent of $Z^S$ and $Z^v = \rho Z^S + \sqrt{1-\rho^2}Z$.
- Generate a Poisson random variable $N_{\Delta t} \sim \ Poisson(\lambda.\Delta t)$ to give the number of jumps on the interval.
- Generate the jumps, where $Y_j \sim \ N(\log(1+\bar{k}) - \frac{\delta^2}{2}\ ,\ \delta^2)$

## 5.2 Stochastic Volatility with Double Exponential Jumps

This model, termed the SVJ-DE model, will assume that the jumps in the underlying asset are distributed with a double exponential distribution (as in the Kou model). This process has risk

neural dynamics given by

$$\frac{dS_t}{S_t} = (r+\omega)dt + \sqrt{v_t}dW_t^S + dZ_t \tag{5.15}$$
$$dv_t = \kappa(\theta - v_t) + \sigma_v\sqrt{v_t}dW_t^v \tag{5.16}$$

the difference between this model and the Bates model lies in the jump component so that

- $W_t^S$ and $W_t^v$ are Brownian motions with correlation $\rho$.
- $Z_t$ is a compound Poisson process with intensity $\lambda$ and distribution of jump sizes given by $V$.
- Jumps are independent of $W^S$ and $W^v$.

The equation for $\log(S_t)$ is given by Itôs formula as

$$dX_t = (r - \frac{1}{2}v_t + \omega)dt + \sqrt{v_t}dW_t^S + d\tilde{Z}_t \tag{5.17}$$

where $d\tilde{Z}_t$ is a compound Poisson process with intensity $\lambda$ and the distribution of jump sizes ($Y = \log(V)$) is given by

$$Y =^d \begin{cases} \xi^+ & \text{with probability } p \\ -\xi^- & \text{with probability } 1-p \end{cases} \tag{5.18}$$
$$\xi^+ \sim \exp\left(\frac{1}{\eta_1}\right) \tag{5.19}$$
$$\xi^- \sim \exp\left(\frac{1}{\eta_2}\right) \tag{5.20}$$

The model parameters can be grouped as follows

| Diffusion | | Jump | |
|---|---|---|---|
| $\kappa$ | speed of mean reversion | $\lambda$ | jump intensity |
| $\theta$ | long variance | $p$ | probability of upward jump |
| $\rho$ | correlation | $\xi^+$ | upward jump parameter |
| $\sigma_v$ | volatility of volatility | $\xi^-$ | downward jump parameter |
| $v_0$ | short variance | | |

**Characteristic Function**

As in the Bates model, the specification of the characteristic function is given by equation 5.1 where $\varphi_{X_t^c}$ (equation 5.2) is the characteristic function of the continuous component and $\varphi_{X_t^j}$ is the characteristic function of the jump component:

$$\varphi_{X_T^j} = \exp\left\{iu\omega T + iu\lambda T\left[\frac{p}{\eta_1 - iu} - \frac{1-p}{\eta_2 + iu}\right]\right\} \tag{5.21}$$

where $\omega$ is chosen by imposing $\varphi_T(-i) = 1$ to ensure the discounted asset price is a martingale, as per section 3.7.3, giving

$$\omega = -\lambda\left[\frac{p}{\eta_1 - 1} - \frac{1-p}{\eta_2 + 1}\right]. \tag{5.22}$$

**Simulation**

Simulation follows that of the Bates model (as in subsection 5.1), except that jumps must be sampled from a double exponential distribution. The volatility diffusion process can be discretized as

$$v_{i+1} = \left(\sqrt{v_i} + \frac{\sigma_v}{2}\sqrt{\Delta t}Z^v\right)^2 + \kappa(\theta - v_i)\Delta t - \frac{\sigma_v^2}{4}\Delta t \tag{5.23}$$

and the logarithm of the asset price can be discretized as

$$\log S_{i+1} = \log S_i + \mu_i^{\mathbb{Q}}\Delta t + \sqrt{v_i \Delta t}\; Z^S + \sum_{j=1}^{N_{\Delta t}} Y_j \tag{5.24}$$

Where the risk-neutral drift term is defined by

$$\mu_i^{\mathbb{Q}} = r - q + \omega - \frac{1}{2}v_i, \tag{5.25}$$

$$\omega = -\lambda\left[\frac{p}{\eta_1 - 1} - \frac{1-p}{\eta_2 + 1}\right] \tag{5.26}$$

Therefore, simulating the asset price over time step $\Delta t$ would involve the following steps:

- Generate two standard normal random variables $Z^S$ and $Z$; where $Z^S$ generates the Brownian motion which drives the asset price process, $Z$ is independent of $Z^S$ and $Z^v = \rho Z^S + \sqrt{1-\rho^2}Z$.
- Generate a Poisson random variable $N_{\Delta t} \sim \; Poisson(\lambda.\Delta t)$ to give the number of jumps on the interval.
- Generate the jumps in the log-price which are sampled from a double exponential distribution, as described in section 4.1.

# Chapter 6

# Market Data and Model Calibration

## 6.1 Data

Market data of option prices was obtained from www.cboe.com, www.optionetics.com and from www.historicaloptiondata.com. This data was cross checked with data from a Bloomberg terminal, to ensure correctness.

Details of the data are summarized in the following table:

| Option Price Data | |
|---|---|
| Trade Day | 22 January 2008 |
| Underlying Asset | S&P 500 Index (ticker symbol SPX) |
| Spot Price of Underlying | US $ 1310.50 |
| Option Prices Included | Bid, Ask and Last Trade Prices |
| Range of Strike Prices | $ 800 to $2000 |
| Strike Price Increments | $ 5 |
| Option Expiry Days | Third Friday of the expiry month |
| Days to Option Expiry | 25, 60, 88, 116, 151, 242, 333, 515, 697, 1061 |
| Risk Free Rate | 3 month US T-Bill rate of 2.27% |
| Assumed Dividend Yield | 2.11% |
| Additional Data | Trade volumes, Open interest |

Mid-Prices of call options, calculated as the average of bid and ask option prices, will be used for analysis and calibration. Mid-prices are often ”better behaved” when compared to last trade prices which may be influenced by trade volumes/liquidity. The call option prices were grouped from strike price levels which had the most information; that is, strike price levels which had listed option prices for at least 7 of the 10 maturities. Thus a data set consisting of 197 call option prices was formed, with strike prices ranging from $1000 to $1900. The option price data used for model calibration and comparison is given in Appendix A.

### 6.1.1 Smoothing the Implied Volatility Surface

The aforementioned option prices are used to extract the instantaneous market implied volatility surface. Implied volatilities are embedded in market option prices, and it is common for market option prices to be quoted directly in terms of implied volatility. The Black-Scholes implied volatility is the value $\sigma_{BS}$ such that for strike price $K$ and maturity $T$:

$$C_{market}(K,T) = C_{BS}(\sigma_{BS}, K, T) \tag{6.1}$$

That is, the Black-Scholes implied volatility is the volatility value which, for strike $K$ and maturity $T$, equates the Black-Scholes option price to the Market option price. The implied volatility surface is a plot of implied volatilities for different strike prices and times to maturity. In order to generate the implied volatility surface, it is necessary to calculate the implied volatility for each option price on the grid. This is done by using numerical methods to find the zero root of the following objective function:

$$f_{BSI}(\sigma_{BS}) = C_{market}(K,T) - C_{BS}(\sigma_{BS}, K, T) \tag{6.2}$$

The Matlab function $fzero$ is used to find the implied volatility which satisfies $f_{BSI}(\sigma_{BS}) = 0$.

The SPX option prices form a grid which may have omissions for some strikes or maturities and, in addition, the implied volatility surface may have irregularities arising from liquidity effects. It is therefore desirable to smooth the surface before comparison between market and model volatilities. Cont and Fonseca (2002) use a non-parametric approach whilst Dumas et al. (1998) and Gatheral (2004) propose a parametric model to fit the implied volatilities.

**Method Used:**

The parametrization of Gatheral (2004) is used and is defined by:

$$\sigma^2_{BS}(k) = a + b\left\{\rho(k-m) + \sqrt{(k-m)^2 + \sigma^2}\right\} \tag{6.3}$$

where $\sigma^2_{BS}$ is the Black-Scholes implied variance and $\sigma_{BS}$ is then the Black-Scholes implied volatility. $k$ = $\log(K/F)$ where $F$ is the forward price for the respective maturity, given by $F_t = S_t e^{(r-q)t}$. $a, b, \sigma, \rho$ and $m$ are model parameters.

Since the model attempts to parameterize the implied volatility function with respect to $\log(K/F)$, which is independent of time, it is necessary to parameterize each time slice separately. Fitting the model to each time slice is achieved by simple optimization (performed by Excel's Solver) of function $\min[\sum_{i=1}^{n}(\sigma_{BSmarket}(k) - \sigma_{BSmodel}(k))^2]^{1/2}$. The Solver optimizer is very sensitive to initial parameter values, hence initial parameter values are approximated as per Vogt (2005). The fits for the Gatheral (2004) Stochastic Volatility Inspired (SVI) parametrization are shown in figure 6.1. The model attempts to fit mid-prices; market bid and ask price implied volatilities are plotted to give additional insight into market implied volatility behavior. Bid price implied volatilities are given by the red crosses and those from ask prices are given by the blue circles. Time (years) to maturity for each set of options is stated on each plot.

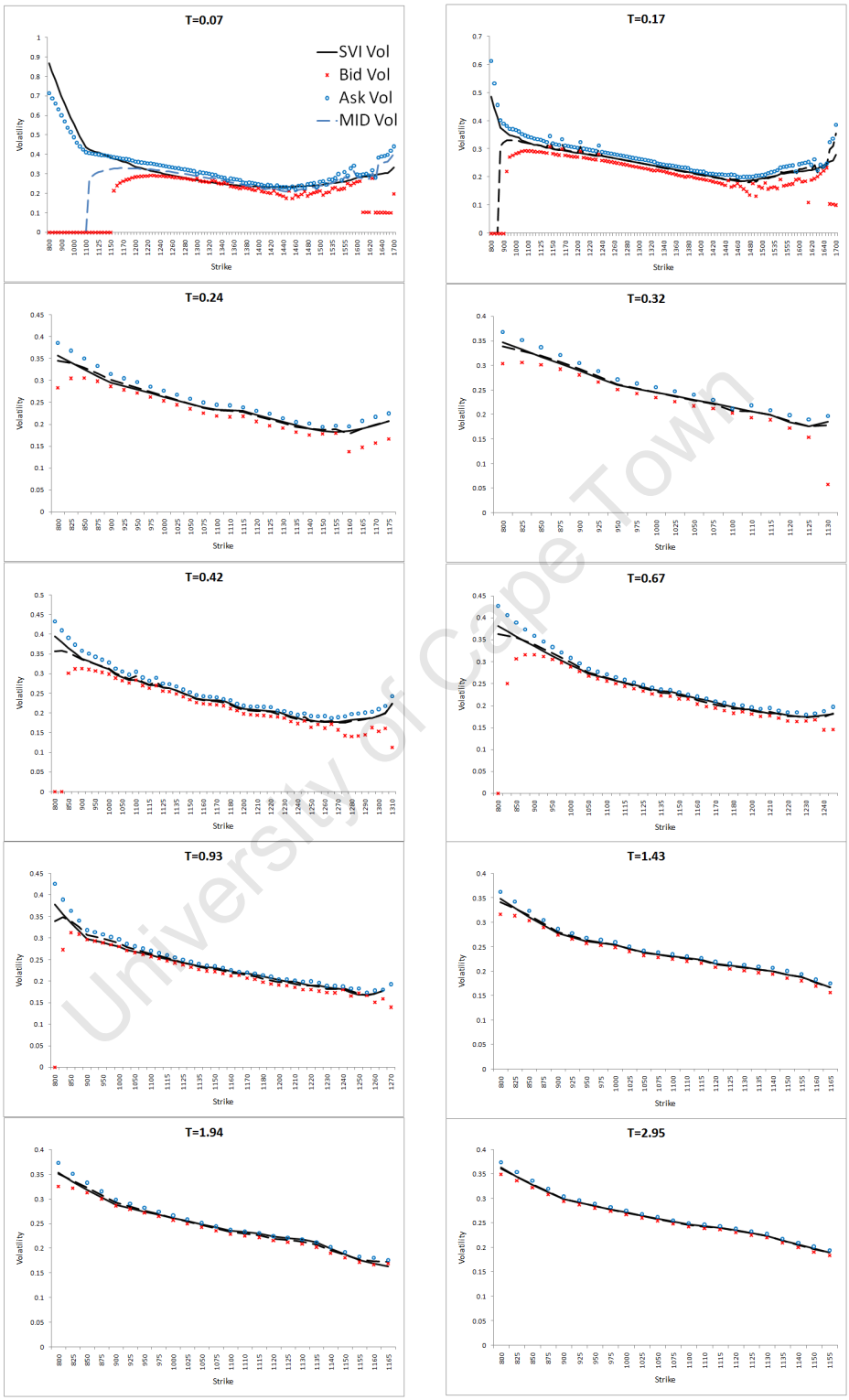





Figure 6.1: SVI parametrization fits to market implied volatility mid-prices.

The following observations are made from the implied volatility curves in figure 6.1.

- The curves for shorter expirations tend to have relatively more of the volatility smile effect, whilst curves for longer expirations show relatively less smile effect but do have the volatility skew effect.
- Implied volatilities from bid and ask prices tend to be less consistent with each other for shorter expirations; and also for deep in-the-money and deep out-of-the-money options.
- Short expiration options can have implied volatilities which are not well behaved, this can be seen by looking at the implied volatilities from the bid prices for options with T=0.07 (time to maturity 25 days).

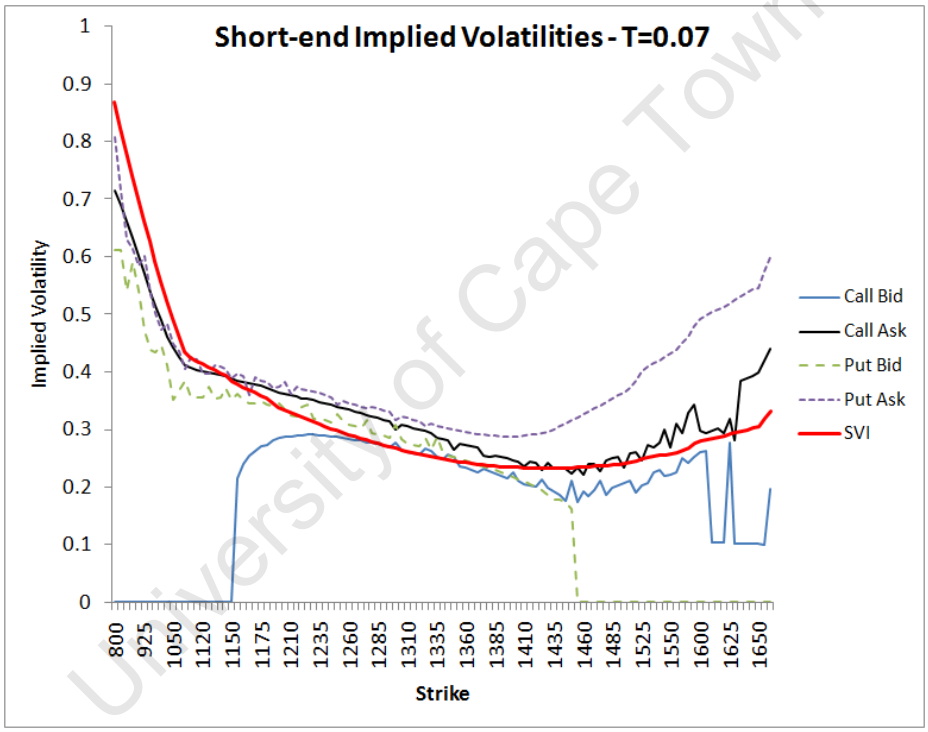





Figure 6.2: Market implied volatility slice.

To further analyze the implied volatility dynamics for shorter expiration options, consider figure 6.2 which shows the implied volatilities extracted from bid and ask prices for both call and put options. Implied volatility behavior is quite inconsistent for deep in-the-money and deep out-of-the-money options (spot price 1310.50).

Finally, the smoothed volatility surface is shown in figure 6.4. The plotted surface shows smile and skew effects over strike prices and times to expiration.

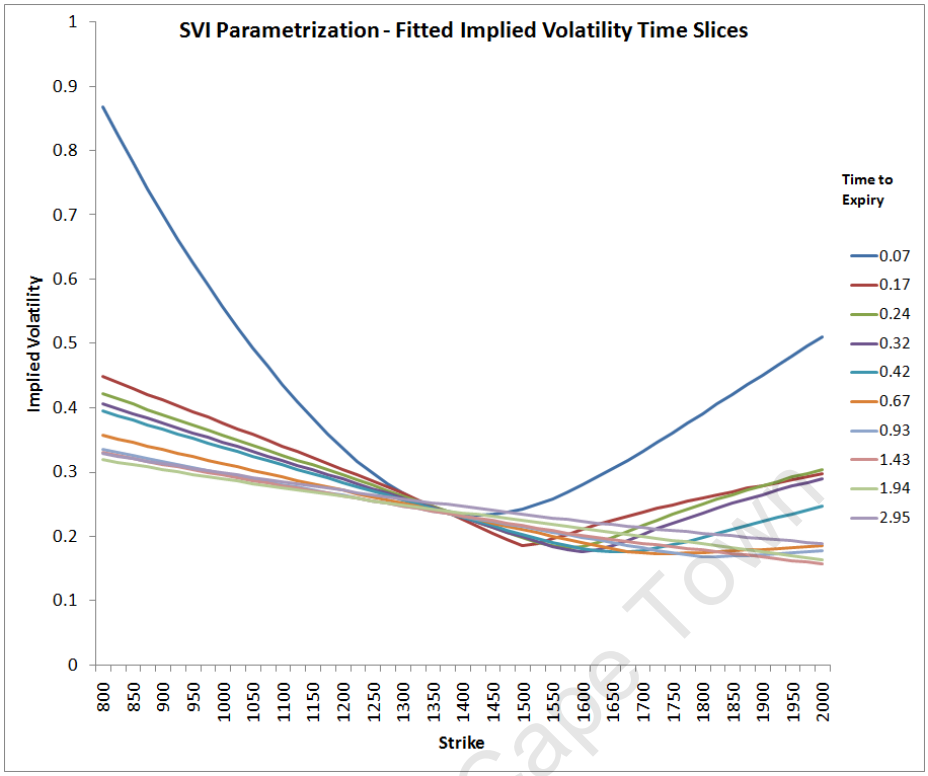



Figure 6.3: Implied volatility curves using fitted parameters.

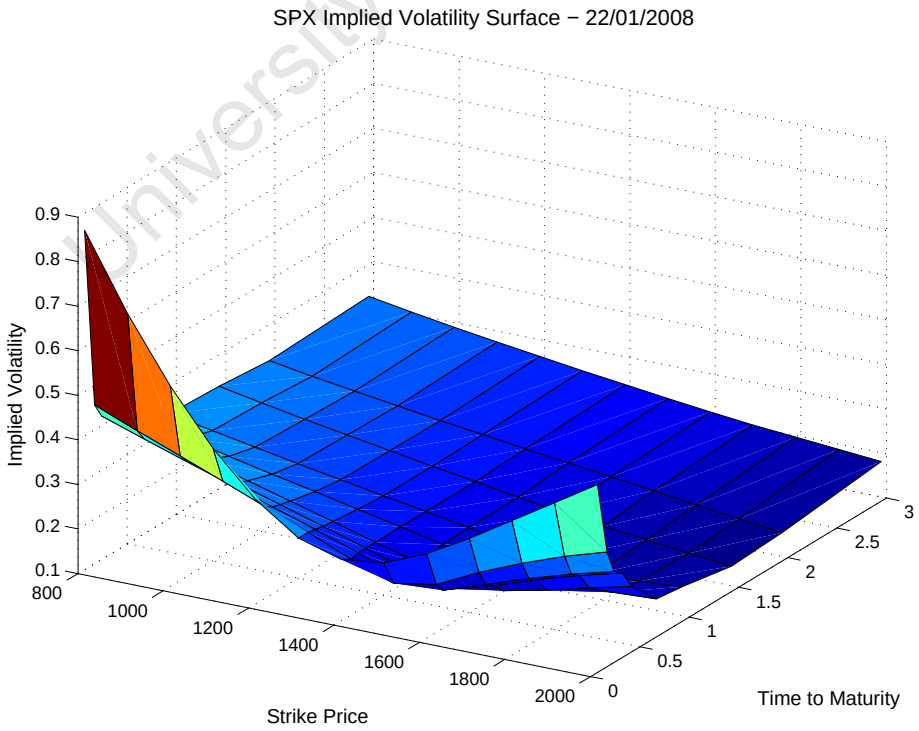



Figure 6.4: Market instantaneous implied volatility surface (SPX Index Options).

## 6.2 Calibration

The calibration process involves fitting an option pricing model to an appropriate set of market prices of vanilla options; this helps to ensure that the option pricing model is pricing exotic over the counter options consistently with the market prices of vanilla options. Model fit is determined by an objective function whose calculation is based on the difference between fitted model prices of vanilla options and current market prices.

The calibration problem is to minimize this objective function, thereby fitting the model to the market. The most commonly used objective function is the squared error between market (vanilla call) option prices and model (vanilla call) option prices:

$$\min_{\vartheta} \sum_{i=1}^{N} \sum_{j=1}^{M} w_{ij} \cdot \Big[ \mathrm{C}_{market}(K_i, \tau_j) - \mathrm{C}_{model}(K_i, \tau_j, \vartheta) \Big]^2 + \mathrm{Penalty}(\vartheta, \vartheta_0) \tag{6.4}$$

where the $K_i$ are the strike prices, the $\tau_i$ are the times to expiry and $\vartheta$ is the model dependent parameter vector, e.g. for the Heston model we would have $\vartheta = (\kappa, \theta, \rho, \sigma_v, v)$. Mikhailov and Nogel (2003) give $\parallel \vartheta - \vartheta_0 \parallel^2$ as an example of a penalty function ($\vartheta_0$ is the initial parameter vector); in this case the penalty function would facilitate model stability by ensuring that a new solution is only chosen if the difference between the new vector and initial vector is offset by the reduction in squared deviations. This ensures that the optimal parameter set will not unnecessarily fluctuate on an intra-day and day-to-day basis. In the context of this dissertation, the aim is to find an initial optimal parameter vector, hence the calibration objective function used is that given in equation 6.4 but with the penalty function omitted. Both unit weights and vega weights were used, as described below.

**Calculation of Model Prices**

The model prices of vanilla options are calculated using the direct integration method as per section 3.7.3. The integral in formula 3.47 is evaluated numerically using a 64-point Gauss-Legendre quadrature.

**Choice of Weights**

The next consideration would be the choice of weight factors, which could lead to differences in calibration results. A relatively large weight is used to convey confidence in a particular price since the error relating to that price will be weighted up; vice versa for a relatively small weight.

Cont and Tankov (2003) suggests the set of weights given by $w_i = \frac{1}{|bid_i - ask_i|}$. In this case, if there is a large bid-ask spread then the respective deviation will be given less weight in the minimization problem, allowing more flexibility for that particular model price. Here the bid-ask spread is used as a proxy for liquidity which is used to convey confidence in option prices.

Cont and Tankov (2003) chapter 13.2 notes that implied volatilities are proportional to bid-ask spreads (for options not too far from the money). This motivates a calibration based on minimizing

the least-square error of Black-Scholes implied volatilities (denoted by $I$).

$$\sum_{i,j} \Big( I(C_{market}(K_i, \tau_j)) - I(C_{model}(K_i, \tau_j, \vartheta)) \Big)^2 \tag{6.5}$$

However, having to numerically compute implied volatilities during calibration is a computational disadvantage which could lead to increased calibration times. Cont and Tankov (2003) therefore uses the following approximation:

$$\sum_{i,j} \left( \frac{\partial I}{\partial C} \Big( I(C_{market}(K_i, \tau_j)) \Big) \Big( C_{market}(K_i, \tau_j) - C_{model}(K_i, \tau_j, \vartheta) \Big) \right)^2 \tag{6.6}$$

and noting the definition of the Black-Scholes vega as the change in option price with respect to volatility leads to

$$\sum_{i,j} \frac{1}{vega^2(I(C_{market}))} \Big( C_{market}(K_i, \tau_j) - C_{model}(K_i, \tau_j, \vartheta) \Big)^2, \tag{6.7}$$

therefore, only the *market* implied volatilities are used in the calibration scheme. Weights that will be considered are defined as

$$w_{ij} \;=\; \min(\frac{1}{vega_{ij}}, 1) \tag{6.8}$$

where vega is computed using the market implied volatility. If an option has a low vega then the option price is less sensitive to changes in volatility, vega is usually lower for options further from the money. Using this form of weight gives less flexibility for option prices which are further from the money, since they have lower vega values. Ensuring that the weights do not exceed 1 (as in Hamida and Cont (2005)) is way of ensuring that too much weight is not given to far from the money options.

Yet another calibration function is given in Schoutens et al. (2003) and Schoutens (2003) who define and use the root-mean-square error objective function to measure fit:

$$\text{RMSE} \;=\; \sqrt{\sum_{i,j} \frac{\big( C_{market}(K_i, \tau_j) - C_{model}(K_i, \tau_j, \vartheta) \big)^2}{C_{market}(K_i, \tau_j)}} \tag{6.9}$$

### 6.2.1 Optimization

Having defined an appropriate objective (cost) function, focus now turns to the task of actually minimizing the objective function. Pricing models can have a number of parameters which need to be optimized; for example, the Heston model has 5 such parameters. Additionally, there may also be constraints on the parameter values. It is common knowledge that the possible objective function or functions may not be well behaved; they may have multiple local minima, therefore, local optimization algorithms may only find a locally optimal solution. These algorithms also tend to require a set of initial parameter values to be specified; and the optimization results can be sensitive to the initial parameter values. On the other hand, if a previously optimal parameter

set is used as input then a local optimizer may produce good and fast results. Local optimizers are appealing due to the fact that they are much faster than other algorithms which attempt to find global minima. The Matlab optimizer *lsqnonlin* was implemented; it was designed to solve nonlinear least-squares (nonlinear data-fitting) problems. However, this optimizer is a local optimizer and for the purposes of initial model fitting it would be mathematically optimal to use a global optimizer. Global optimization methods use algorithms which attempt to find a globally optimal solution; hence they tend to be much slower than local optimizers. A comparison of optimization methods for the Heston model can be found in Moodley (2005).

The Adaptive Simulated Annealing (ASA) method was the primary optimizer used for model calibration and is a global optimization method. The c-language code is open-source and is downloadable from www.ingber.com, which also contains documentation and papers on ASA. ASAMIN is a gateway to the ASA program which allows it to be used with Matlab; it was developed by Shinichi Sakata and details on implementation can be found in Moins (2002). The ASA algorithm has been applied to a variety of problems in a number of different areas including neuroscience and finance. The algorithm was designed to fit empirical data to a theoretical cost function with a D-dimensional parameter vector with or without constraints. The algorithm starts with a so called high *temperature*, which can be thought of as the energy in a particle. The more energy, the farther the particle can move, and similarly, the value of the parameter vector can be changed with more freedom. Generating distributions are then used to define where in the parameter space the method will search. With a high temperature and defined generating distributions, the method can explore a large space and identify multiple local minima. An acceptance distribution comes into play by allowing the method to move, with a given probability, to a less optimal parameter set. This allows the process to move amongst (and not become trapped in) local minima. The generating distributions and acceptance distributions depend on the temperature, which is reduced as the method progresses. A example of another global optimization method is one based on the *Differential Evolution* algorithm.

It is desirable to consider the optimization problem from a more practical perspective, say, from the perspective of an exotic equity option dealer, or a major player in exotic equity options such as an investment bank. These institutions have well developed information technology systems (including grid computing systems) which allow for real-time pricing and real-time risk monitoring. Now, as the market changes, the prices of vanilla options change and hence then price of exotic options should also change accordingly. However, ensuring that the real-time price of exotic options is consistent with prices of vanilla options would require real-time model calibration. This is where the differences in local and global optimization methods become emphasized due to the differences in computational time. One common solution would be to use a local optimization method intra-day, when parameter values are unlikely to fluctuate much under normal market conditions. A global optimization method can then be used at the end of each trading day or when mark-to-market takes place. Additionally, under extreme market conditions such as market crashes, global optimization methods may be activated as parameter vectors would be likely to change significantly.

# Chapter 7

# Model Comparison

## 7.1 Calibration Results

Table 7.1 shows the results from model calibration; the results are in the form of least-squared error (LSE) values. Calibration was performed as per section 6.2 using equation 6.4 with no penalty function and using the ASA optimizer. Two different choices of weights were used - unit weights ($w_{ij} = 1$), and also vega weights (as defined in equation 6.8). Note that the more parameters that a model has, the longer it may take to calibrate and the more vulnerable the model may be to parameter instability.

Table 7.1: Minimized objective function values (Least Square Errors)

| Type | Model | Unit weights | Vega weights |
|---|---|---|---|
| SV-Diffusion | **Heston** | 814.83 | 233.53 |
| Jump-Diffusion | **Merton** | 9078.72 | 2832.59 |
| Jump-Diffusion | **Kou** | 8725.66 | 1808.07 |
| Pure Jump | **Variance-Gamma** | 9398.96 | 2829.13 |
| Pure Jump | **Variance-Gamma with CIR Time** (VG-CIR) | 4319.88 | 1036.39 |
| SVJ | **Bates** | 545.23 | 186.00 |
| SVJ | **SV + Double Exponential Jumps** (SVJ-DE) | 542.68 | 184.78 |

**Unit Weights**

For the case where unit weights were used, a better fit is determined equally by every option price, no matter how far from the money it is. The resulting LSE's give a clear indication of how the models fit the market, relative to eachother.

- The Merton, Kou and VG models do not give good performance; they all perform similarly with relatively high LSE values between 8726 and 9399. These models have 4,5 and 3 parameters respectively.
- Adding stochastic effects to the VG model to arrive at the VG-CIR model results in a considerable improvement in fit - a reduction in the LSE from 9399 to 4920. Note that this model has 6 parameters.
- The Heston SV model gives a LSE of 815, a considerable improvement over the aforementioned models. The Heston model requires estimation of 5 parameters (1 less than VG-CIR, and with a better fit). In the context of this data set, given the number of parameters and the LSE, the Heston stochastic volatility model outperforms the pure jump and jump diffusion models.
- The SVJ models give the best performance. Compared to the Heston model, the Bates model with its log-normal jumps results in a decrease in the LSE from 815 to 545. The Bates model has a total of 8 parameters, 3 more than the Heston model; so that the reduction in LSE comes with the cost of additional parameters.
- The SVJ-DE model combines double exponentially distributed jumps with Heston-type stochastic volatility and has 9 parameters. The result is a reduction in the LSE to 543 (a 33% reduction compared to the Heston model), the SVJ-DE model therefore gives a substantially better fit than the Heston model and overall gives the best fit (which is slightly better than that of the Bates model).

**Vega Weights**

Where vega weights are used, emphasis is placed on options with low vega values, and these are the options that lie away from the money. The resulting LSE's then give an indication of how the models fit the options which lie further out of the money (OTM) or further in the money (ITM), with a threshold level to ensure too much weight is not placed on options far from the money.

- The Merton and VG models result in the worst fits, with LSE's of 2833 and 2929, respectively.
- The Kou model gives a better fit with an LSE of 1808. Whilst both the Kou and the Merton models are jump diffusion models, the LSE's imply that the double exponential jumps in the Kou model result in a better fit than the log-normal jumps of the Merton model.
- The VG-CIR model gives better performance with an LSE of 1037.
- The SV and SVJ models offer much reduced LSE's. Heston gives an LSE of 234, and the SVJ models offer a 20% reduction of this to 186 for the Bates model and 185 for the SVJ-DE model. In this case, the reduction in LSE would justify the use of a model with more parameters.

Whilst the SVJ models do result in the best fit, there are a number of practical considerations/implications due to the fact that the models have 8 or 9 parameters (more than any other model being considered). Firstly, the larger number of parameters in an SVJ model may lead to parameter spaces with many local minima, so that global optimization methods should be primarily used. Also, parameter vectors may be vulnerable to unnecessary fluctuations, in which case, a penalty function should be used when performing model calibration (as mentioned in section 6.2). Another implication would be increased calibration times; these could be reduced using more efficient calibration algorithms and imposing appropriate limits on the parameter values. Finally, there is the concern of an increase in the time taken to price exotic options via Monte Carlo methods. The Heston model requires simulation of two diffusion processes. The Bates model would additionally require simulation of the jumps from a normal distribution and the SVJ-DE model would require simulation from a double exponential distribution.

### 7.1.1 Model Pricing Error

This section presents plots with model prices and with market prices for the case of unit weights. A perfect fit would result in all of the model prices (represented by + signs) lining up with all of the market prices (represented by circles). Prices from the VG-CIR model are shown in figure 7.1; note that the prices do not line up very well, with noticeable differences. This is the expected due to the relatively high LSE value of 4320.

Heston prices are then shown in figure 7.2; prices generally line up better, this is in accordance with the lower LSE of 815.

Finally, prices from the model with the best fit (the SVJ-DE) model are shown in figure 7.3. Close inspection reveals that prices generally line up better than in the Heston model. These plots serve to visually reinforce the results in the previous section. More insight into model behavior is gained in the following section which focusses on how the implied volatilities of the fitted models compare to the market implied volatilities.

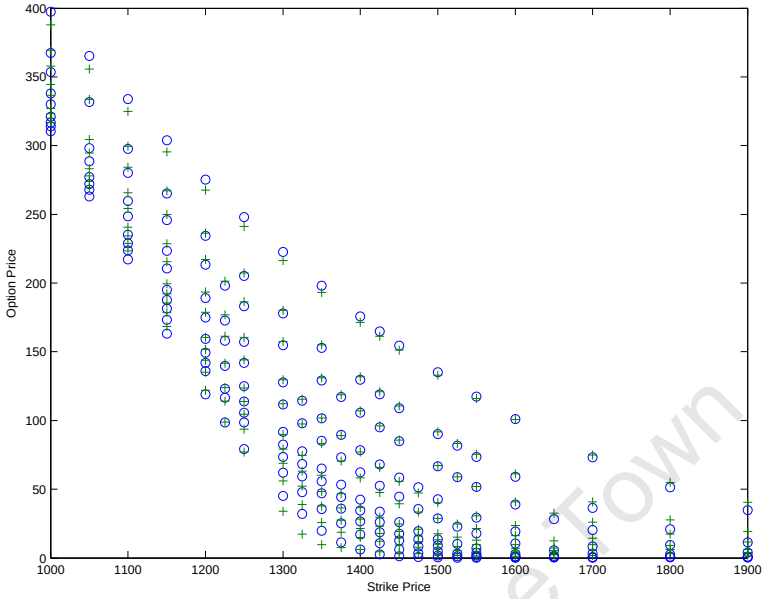



Figure 7.1: VG-CIR prices (+) versus Market prices (o).

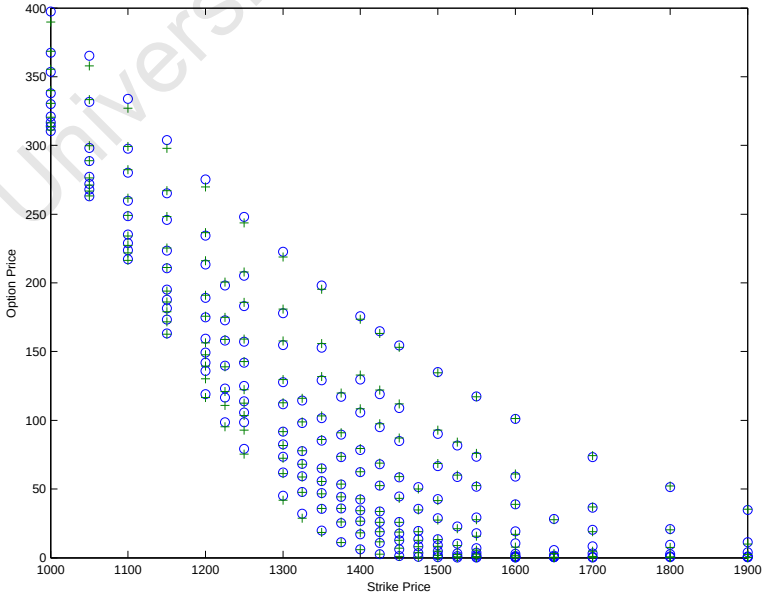



Figure 7.2: Heston prices (+) versus Market prices (o).

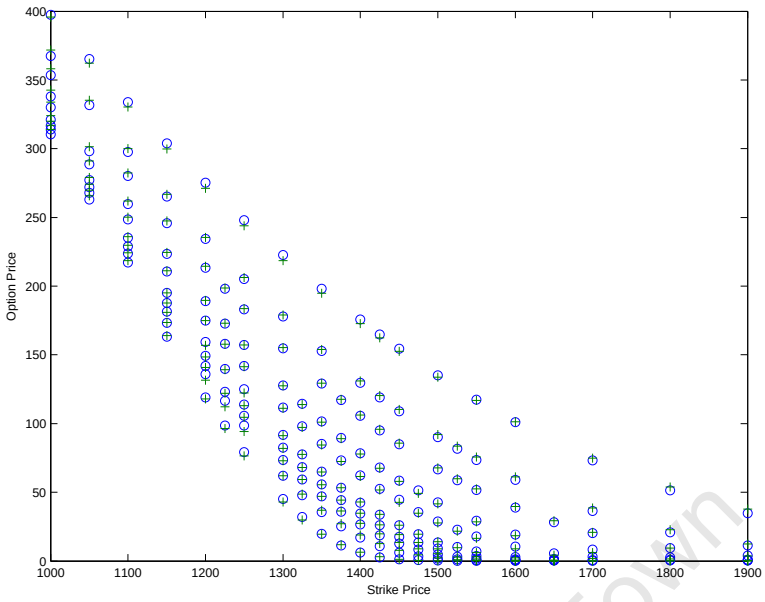



Figure 7.3: SVJ-DE prices (+) versus Market prices (o).

## 7.2 Volatility Surfaces

Insight into model performance and dynamics can be gained by considering model implied volatility surfaces. The benchmark surface is the market (SPX) implied volatility surface (figure 7.4). Note that the volatility smile effect is more prevalent for shorter maturities and the volatility skew effect is more prevalent for longer maturities. A good model should be able to capture these effects. In the previous section, the word *fit* was used to convey how the model prices matched the market prices. In this section, *fit* will be used to convey how the model implied volatilities match the market implied volatilities. The case of unit weights will be considered in this section. Volatility surfaces are shown for each calibrated model, these can be used to give a general idea of how the calibrated models are able to match market volatilities by comparing to the benchmark surface. The implied volatility time slices are also shown and enable a closer comparison to be made since they show model implied volatilities versus market implied volatilities separately for each maturity. The errors (model implied volatility minus market implied volatility) are included on the latter set of figures and give a clearer picture of where the models perform poorly (i.e. for which maturities and strikes).

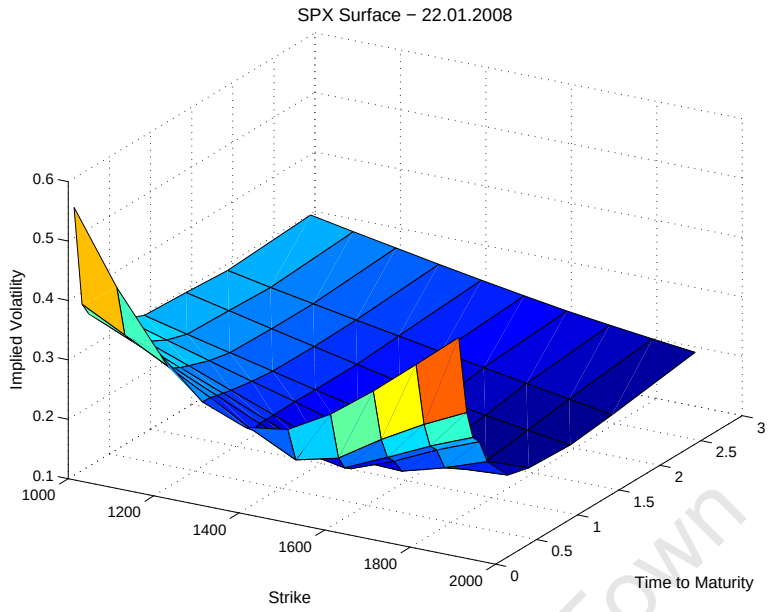



Figure 7.4: SPX instantaneous implied volatility surface.

### 7.2.1 Heston Surface

The first and most noticeable observation from the Heston surface in figure 7.5 is the poor performance on the short end. This can be viewed more clearly by considering the implied volatility time slices and in figure 7.6, noticing the following:

- The T=0.07 slice shows a particularly bad fit. The model is not generating enough smile effect, especially for deep in the money call options where the error drops down to -0.2.
- The fit improves significantly for the next few maturities (T=0.17 to T=0.93), however, it is a general observation that for shorter maturities the model volatilities lie below the market volatilities. The model is unable to adequately capture the volatility smile effect.
- For T=1.43 onwards, good fits are observed since we have errors which are almost zero. This is the portion of the surface where volatility skew effects are more prevalent, so that the model is able to adequately capture volatility skew effects.

Table 7.2: Optimal parameters - Heston model

| | **Parameter** | **Optimal Value** |
|---|---|---|
| rate of mean reversion | $\kappa$ | 1.3883 |
| long variance | $\theta$ | 0.0827 |
| correlation | $\rho$ | -0.9441 |
| volatility of volatility | $\sigma$ | 0.4792 |
| short variance | $v_0$ | 0.0690 |

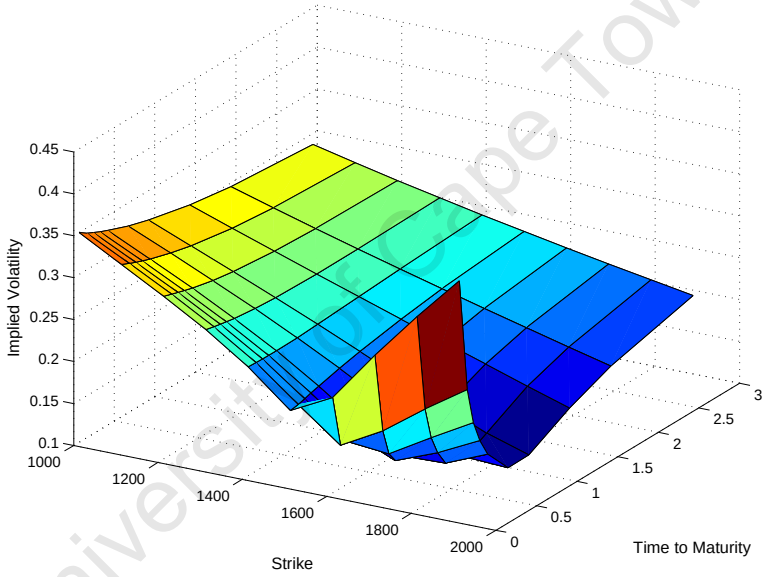



Figure 7.5: Heston implied volatility surface.

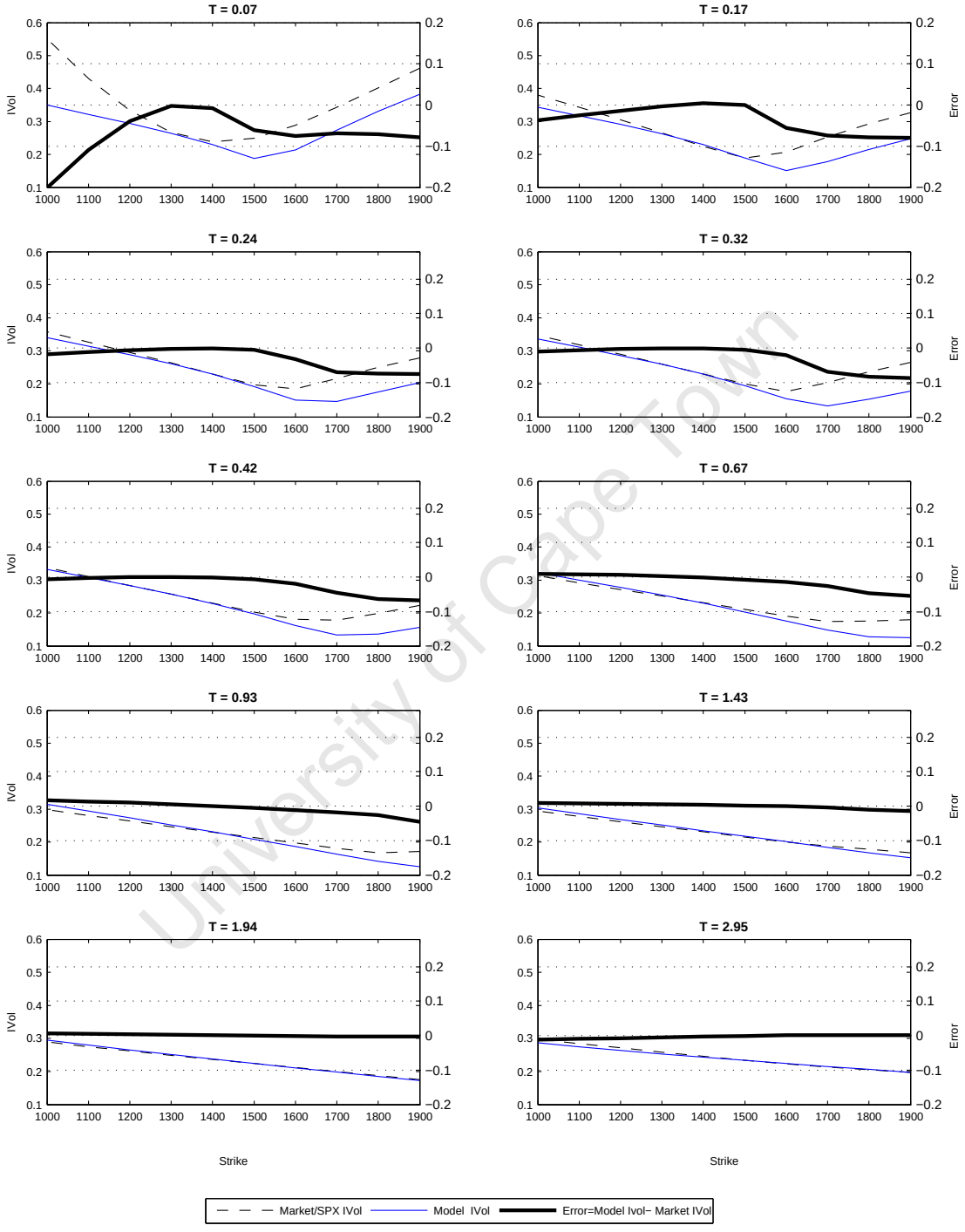





Figure 7.6: Implied volatility time slices: Heston Model. right y-axis = errors, left y-axis = implied volatility

Table 7.3: Optimal parameters - Merton model

| | Parameter | Optimal Value |
|---|---|---|
| volatility of Brownian motion | $\sigma$ | 0.1657 |
| jump intensity | $\lambda$ | 0.2014 |
| jump size distribution mean | $\bar{k}$ | -0.4824 |
| jump size distribution standard deviation | $\delta$ | 0.2381 |

Table 7.4: Optimal parameters - Kou model

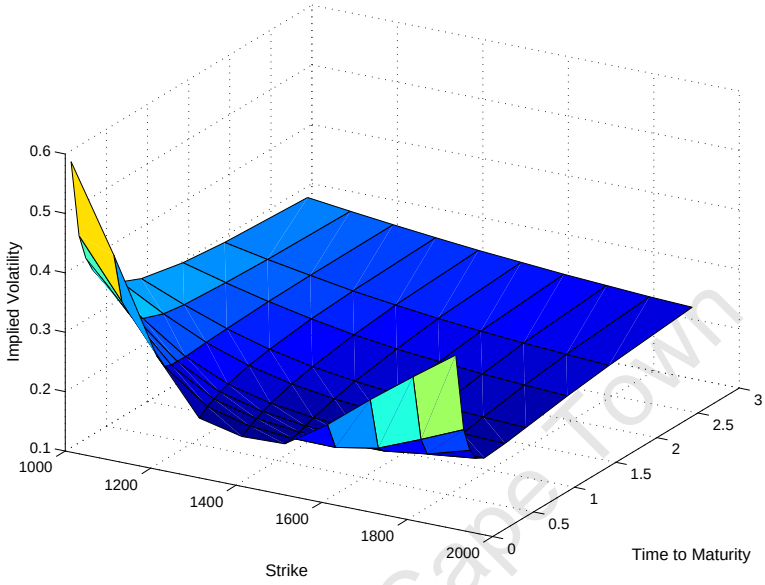
| | Parameter | Optimal Value |
|---|---|---|
| volatility of Brownian motion | $\sigma$ | 0.1452 |
| jump intensity | $\lambda$ | 0.7625 |
| probability of upward jump | $p$ | 0.0000 |
| exponential distribution parameter - upward jumps | $\eta_1$ | 30.0000 |
| exponential distribution parameter - downward jumps | $\eta_2$ | 4.9222 |

### 7.2.2 Jump Diffusion Surfaces

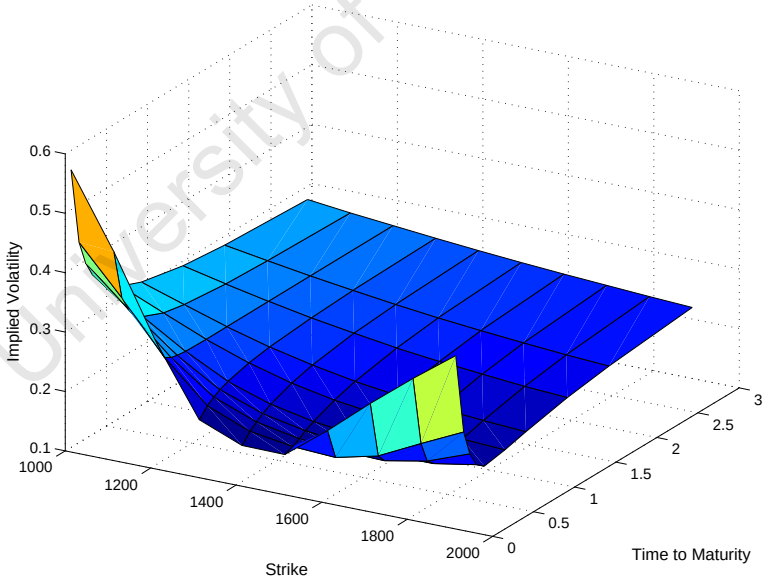
At first glance, figures 7.7(a) and 7.7(b) indicate that the Kou and Merton surfaces seem to exhibit behavior opposite to that of the Heston model, a potentially better fit at the short end (capturing smile dynamics better), with a worse fit at the long end (where the skew effect prevails). The volatility time slices of the Kou model are shown in figure 7.9 and the following observations can be made

- At the short end (T=0.07) the model is able to capture smile effects well, though there is still a noticeable difference between model and market volatilities. The errors of deep in the money call options are much lower than those of the Heston model whilst errors of deep out of the money call options are similar.
- Moving on to longer maturities such as T=0.93, the model gives a good fit.
- At the long end (T=2.95) the model gives a relatively poor fit, this is because the model volatilities are too flat. This shows that the model is unable to generate enough of a skew effect for long maturities (in this case for maturities beyond T=0.93).

To summarize, whilst the Kou model is able to capture smile effects at the short end, it is unable to adequately capture skew effects at the long end. The Merton model is also able to capture smile effects at the short end, however, the Merton model is expected to be less efficient than the Kou model based on the LSE values in the previous section. Note that the Kou model uses different distributions for upward and for downward jumps; from comparison of figures 7.8 and 7.9, this seems to give the model a little more freedom on the short end leading to a better fit for deep in the money call options (when compared to the Merton model).




(a) Merton implied volatility surface: log-normal jumps.




(b) Kou implied volatility surface: double exponential jumps.

Figure 7.7: Jump-diffusion Surfaces

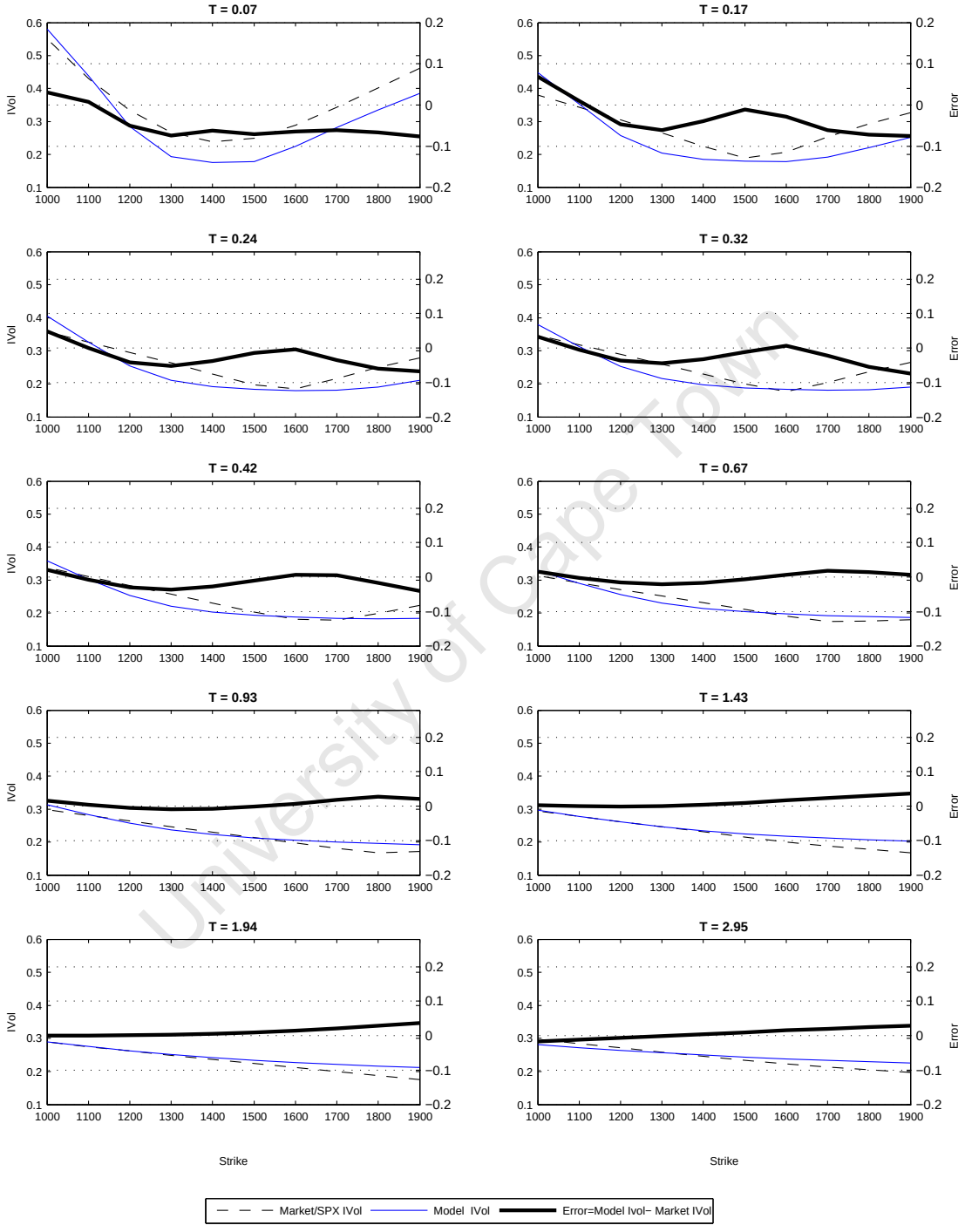





Figure 7.8: Implied volatility time slices: Merton Model. right y-axis = errors, left y-axis = implied volatility

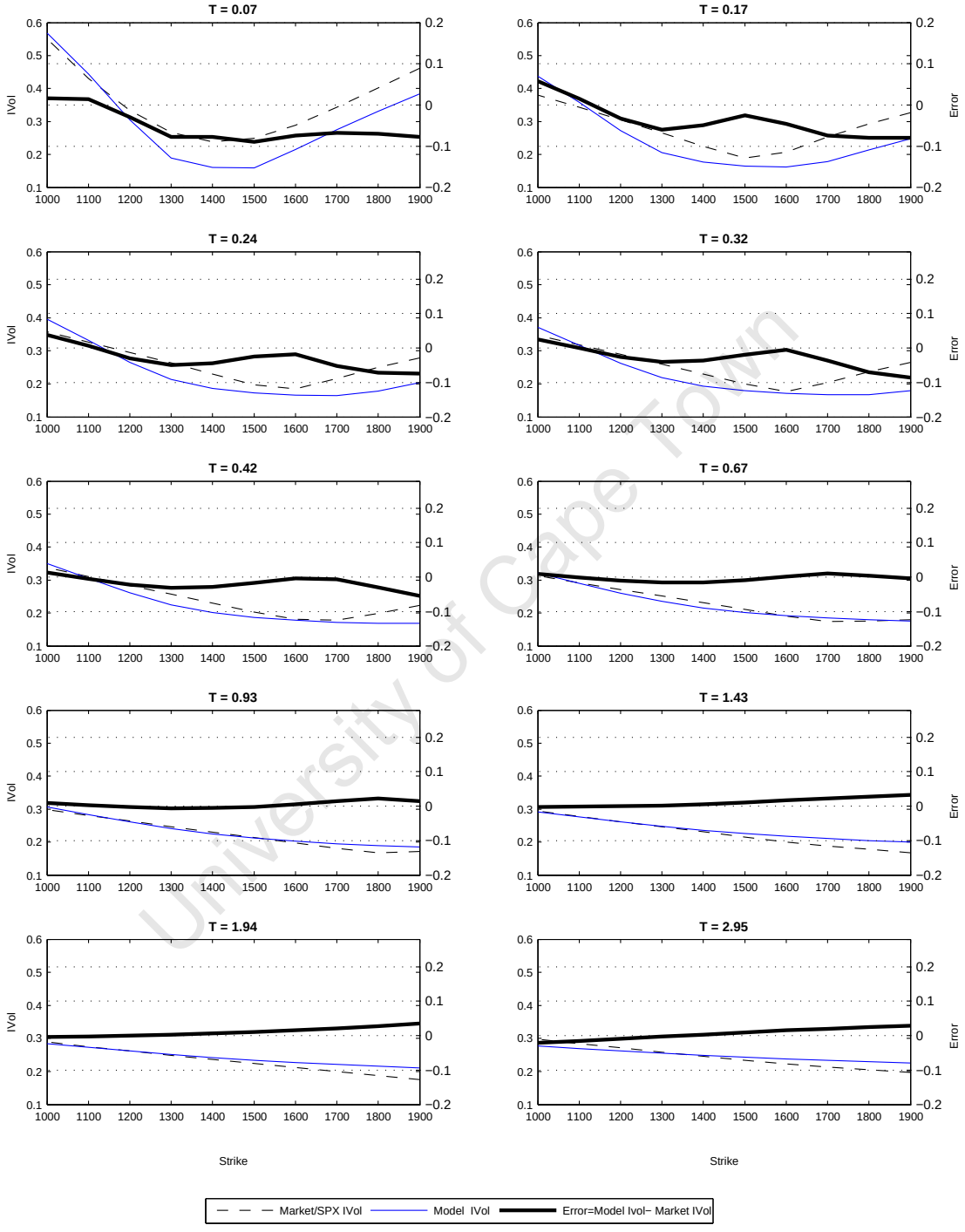





Figure 7.9: Implied volatility time slices: Kou Model. right y-axis = errors, left y-axis = implied volatility

### 7.2.3 Pure Jump Surfaces

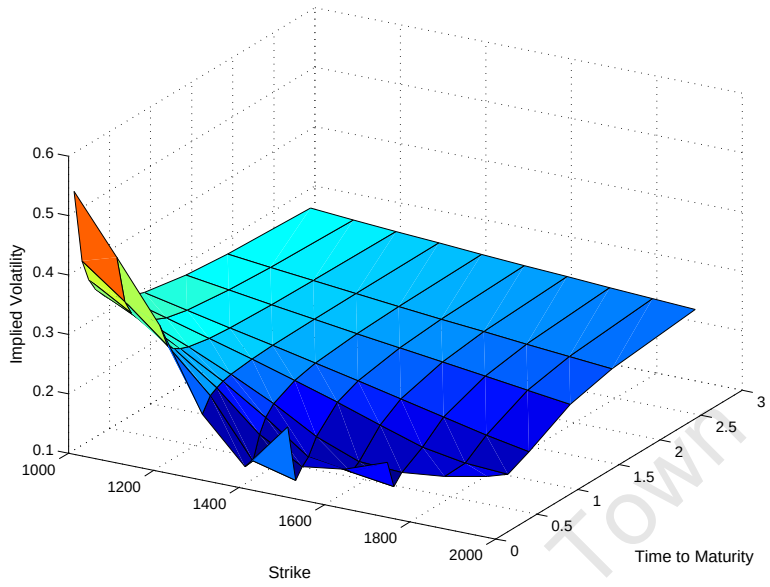



Figure 7.10: VG implied volatility surface.

The volatility time slices from the VG and VG-CIR models (shown in figures 7.12 and 7.13 respectively) show similar patterns to those of the Merton and Kou models. That is, the pure jump models are able to generate adequate smile effects on the short end, but unable to generate enough skew on the long end. To be more specific, consider the VG-CIR model in 7.13:

- At the short end (T=0.07 to T=0.24), the model is more than able to generate smile effects. The problem is that the model is unable to adequately match model volatilities with market volatilities, as seen by the errors.
- T=1.43 and T=1.94 show that the model is able to produce enough of a skew for these maturities and therefore give decent fits (this is in contrast to the VG model as seen in figure 7.12).
- At the last maturity, T=2.95, the model is not quite able to produce enough skew to match the market, but it does perform notably better at this than the other models considered thus far.

Therefore, the results show that pure jump models are not able to adequately capture market dynamics.

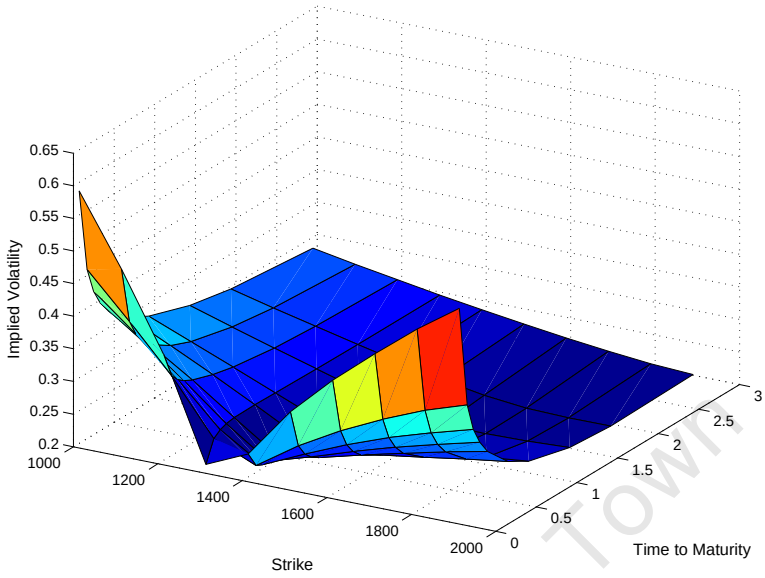



Figure 7.11: VG-CIR implied volatility surface.



Table 7.5: Optimal parameters - VG model

| | **Parameter** | **Optimal Value** |
|---|---|---|
| volatility of Brownian motion | $\sigma$ | 0.1425 |
| parameter of Gamma process | $\nu$ | 0.3264 |
| drift of Brownian motion | $\theta$ | -0.4046 |

Table 7.6: Optimal parameters - VG-CIR model

| | **Parameter** | **Optimal Value** |
|---|---|---|
| | $C$ | 1.4244 |
| | $G$ | 4.5645 |
| | $M$ | 10.0000 |
| rate of mean reversion | $\kappa$ | 2.6466 |
| long-run rate of time change | $\eta$ | 0.7728 |
| volatility of time change | $\lambda$ | 3.0620 |

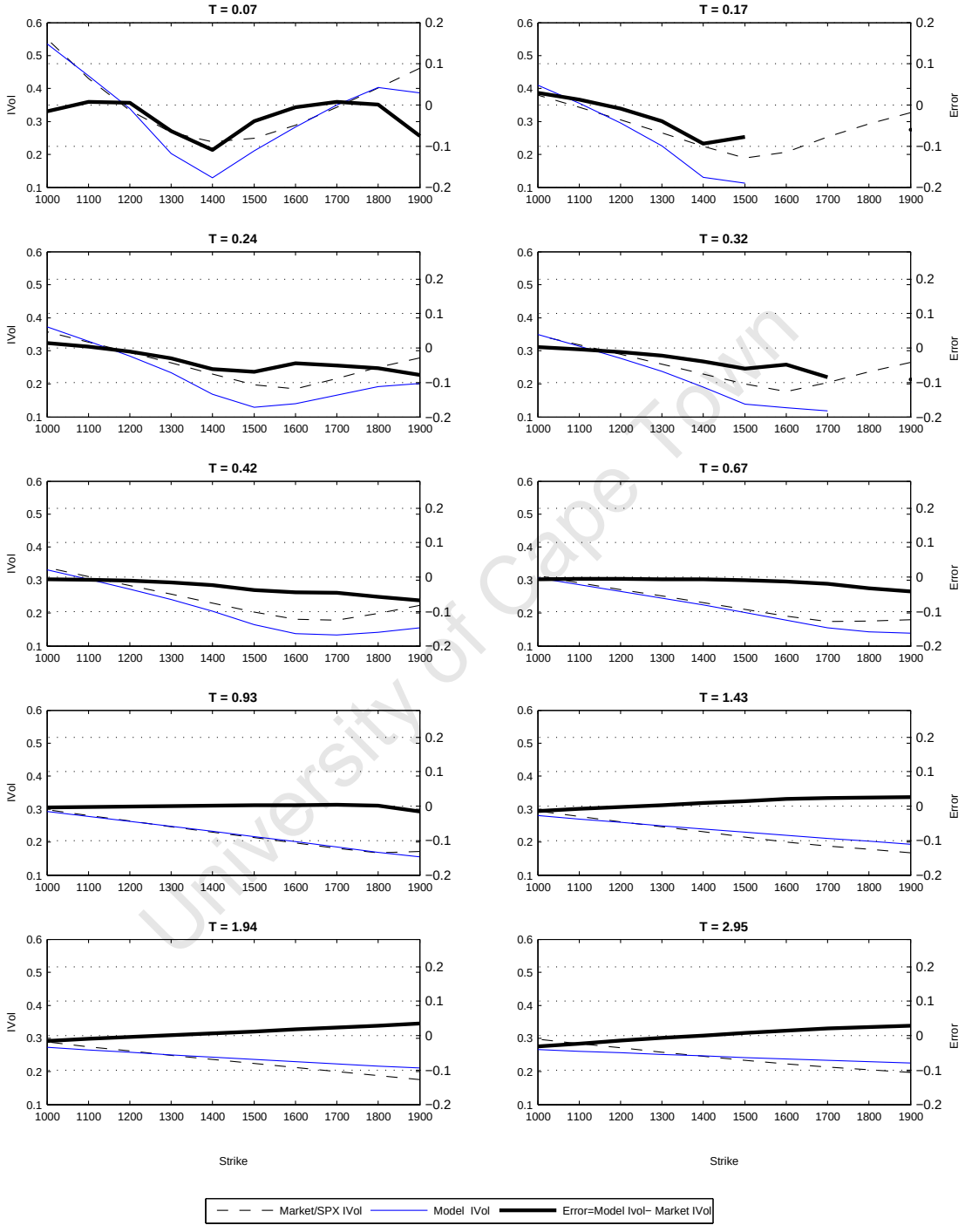





Figure 7.12: Implied volatility time slices: VG Model. right y-axis = errors, left y-axis = implied volatility

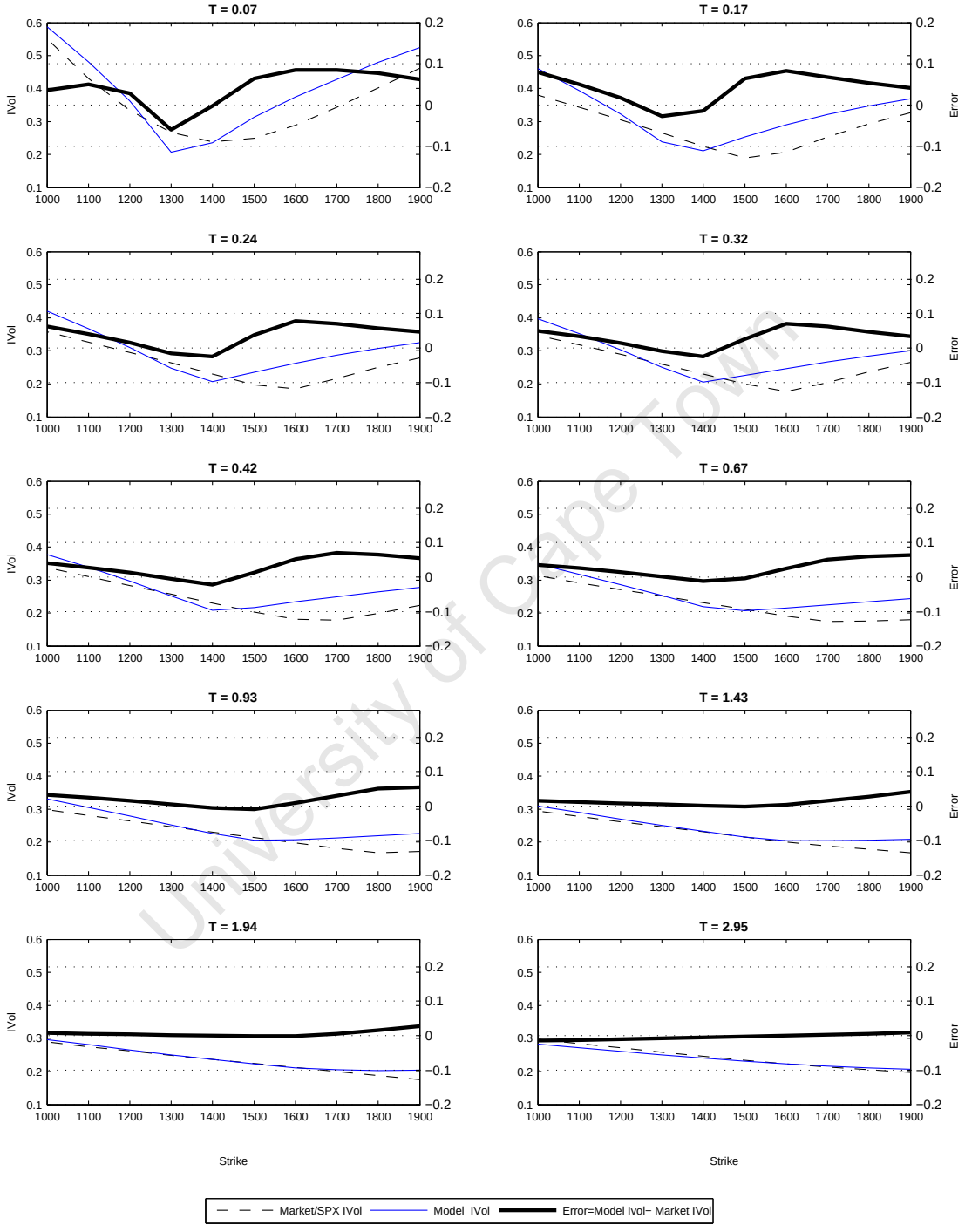



Figure 7.13: Implied volatility time slices: VG-CIR Model. right y-axis = errors, left y-axis = implied volatility

Table 7.7: Optimal parameters - Bates model

| | Parameter | Optimal Value |
|---|---|---|
| rate of mean reversion | $\kappa$ | 3.0533 |
| long variance | $\theta$ | 0.0481 |
| correlation | $\rho$ | -0.8992 |
| volatility of volatility | $\sigma$ | 0.4809 |
| short variance | $v_0$ | 0.0728 |
| jump intensity | $\lambda$ | 0.0176 |
| jump size distribution mean | $\bar{k}$ | -0.9900 |
| jump size distribution standard deviation | $\delta$ | 3.0157 |

Table 7.8: Optimal parameters - SVJ-DE model

| | Parameter | Optimal Value |
|---|---|---|
| rate of mean reversion | $\kappa$ | 3.2723 |
| long variance | $\theta$ | 0.0458 |
| correlation | $\rho$ | -0.9814 |
| volatility of volatility | $\sigma$ | 0.4304 |
| short variance | $v_0$ | 0.0728 |
| Jump intensity | $\lambda$ | 0.0441 |
| probability of upward jump | $p$ | 0.5753 |
| exponential distribution parameter - upward jumps | $\eta_1$ | 15.1256 |
| exponential distribution parameter - downward jumps | $\eta_2$ | 0.0000 |



### 7.2.4 SVJ Surfaces

The SVJ models incorporate both stochastic volatility and jumps in the underlying. The information gained thus far has given the impression that SV models and jump models almost have opposite behavior - the former class fits long maturities better whilst the latter class fits shorter maturities better. SVJ models are an attempt to exploit this behavior and improve model fit by combining the strengths of the two classes. The dynamics of their implied volatility surfaces give more insight, comparing figures 7.16 and 7.17:

- The Bates SVJ model and the SVJ-DE model give the best results when model volatilities are compared to market volatilities. Their results show good fit on the short end (ability to capture smile dynamics), and the long end errors are almost zero which indicates that the models are able to adequately capture skew dynamics.

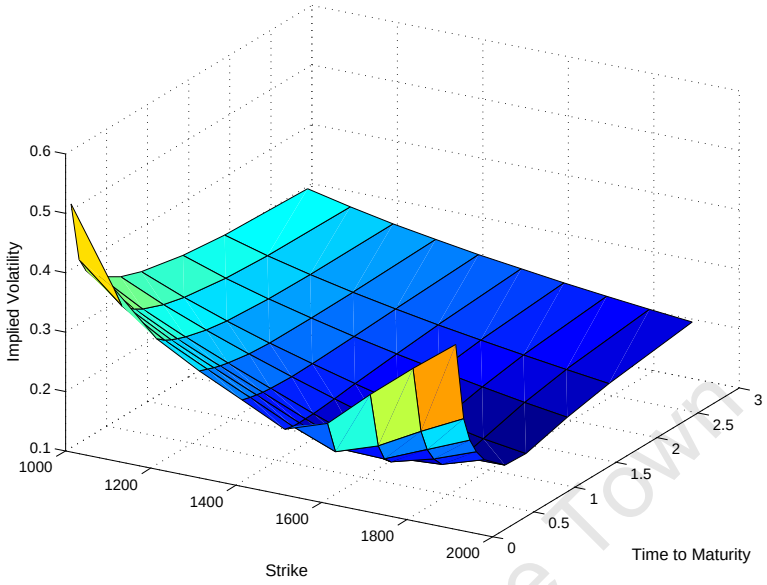



Figure 7.14: Bates implied volatility surface.

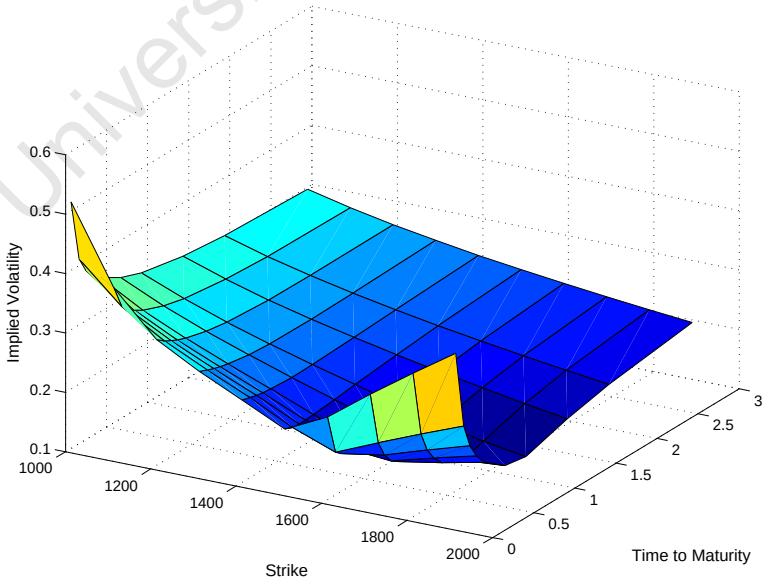



Figure 7.15: SVJ-DE implied volatility surface.

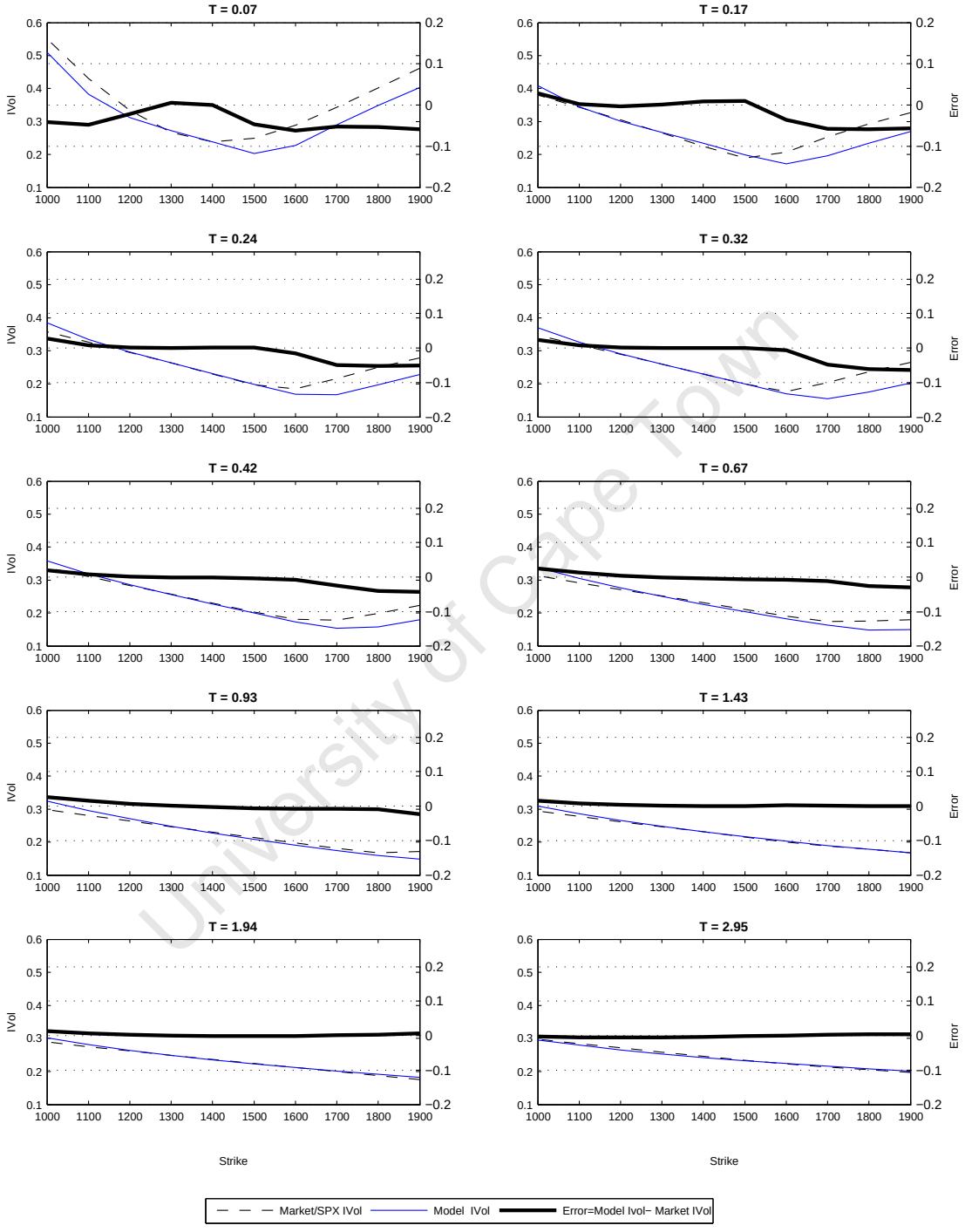



Figure 7.16: Implied volatility time slices: Bates Model. right y-axis = errors, left y-axis = implied volatility

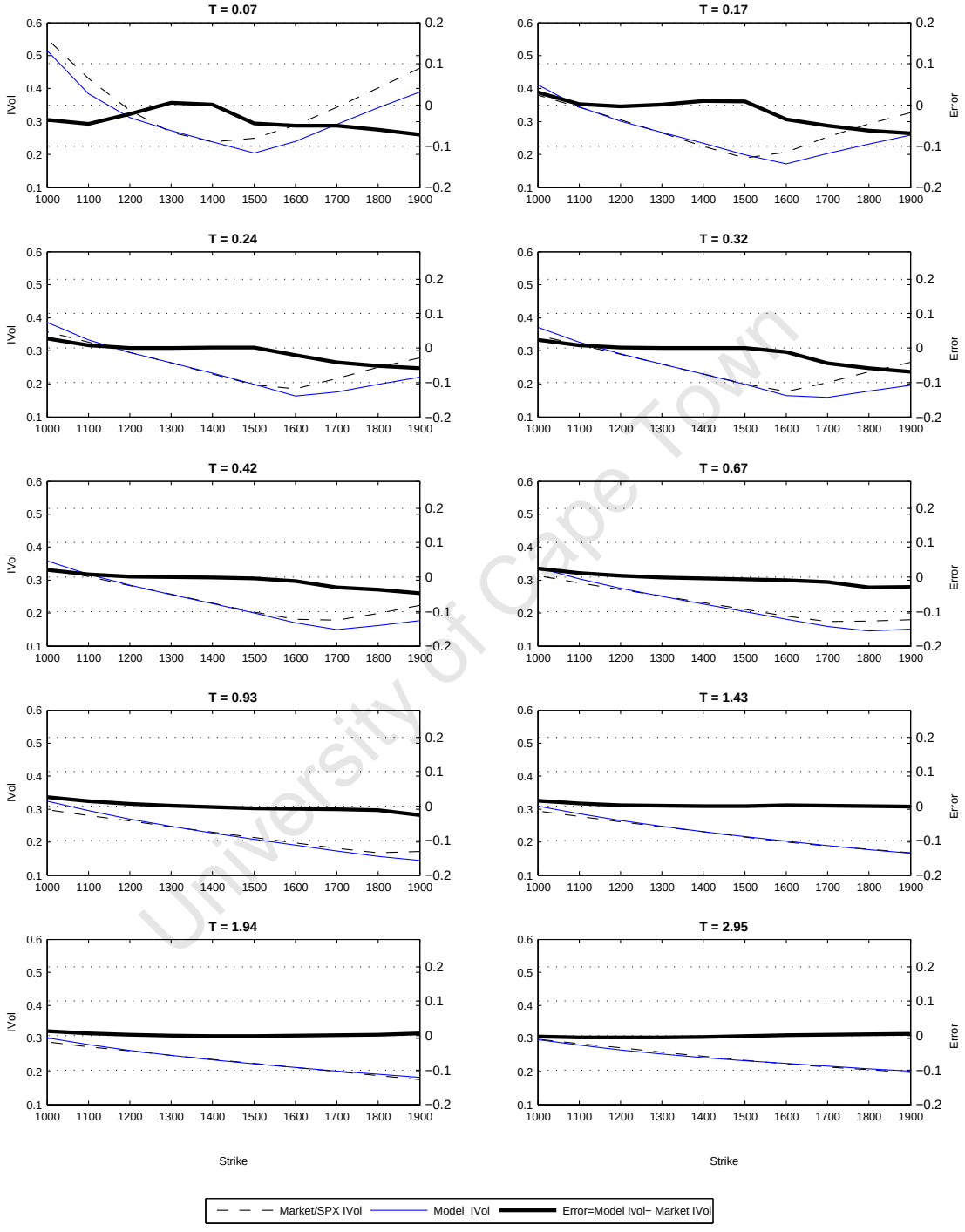



Figure 7.17: Implied volatility time slices: SVJ-DE Model. right y-axis = errors, left y-axis = implied volatility

### 7.2.5 Short Term Model Fit

Finally, figure 7.18 groups the time slices for the shortest maturity (T=0.07) and for all the models to facilitate a comparison of the short term model fit.

- Notice that if we look at volatilities of deep in the money call options (strike $< 1200$) in isolation then we find a weakness of the SVJ models. Begin by recalling that the Heston model is unable to generate enough volatility for these options. Then the jump models are able to generate enough volatility and actually generate too much, giving positive errors. The SVJ models seem to be able to generate more volatility for these options than the Heston model (attributed to the inclusion of jumps), but still not as much as the models without stochastic volatility. The Variance Gamma model gives very good performance for deep in the money call options.

- The models with stochastic volatility (Heston, Bates and SVJ-DE) give the best results for at the money options (Strike =1310). All the other models cannot generate enough volatility for at the money options and hence exhibit negative errors; the VG-CIR model performs best amongst them.

- Moving on to deep out of the money call option strikes, we find an underestimation of volatility (negative errors) for all but the VG-CIR model. SVJ models have smaller errors than Heston, Merton, Kou and VG-CIR. The VG model gives the SVJ models competition for these strikes and the VG model only performs worse at the last strike (1900).

So, for the case of short term fits we see the best results from the SVJ models and from the VG model. The VG model gives a good fit for strikes which are far from the money and the SVJ models give the best fit for strikes which are close to the money.

The final view, based on implied volatility fits to all maturities and also based on LSE values, is that the class of jump diffusion models with stochastic volatility and jumps in the underlying (SVJ models) are found to give the best overall fit. They were the most efficient at capturing overall market implied volatility dynamics and the calibration results in section 7.1 showed that SVJ models produced the lowest LSE values, hence fitted the market data better than the other models.

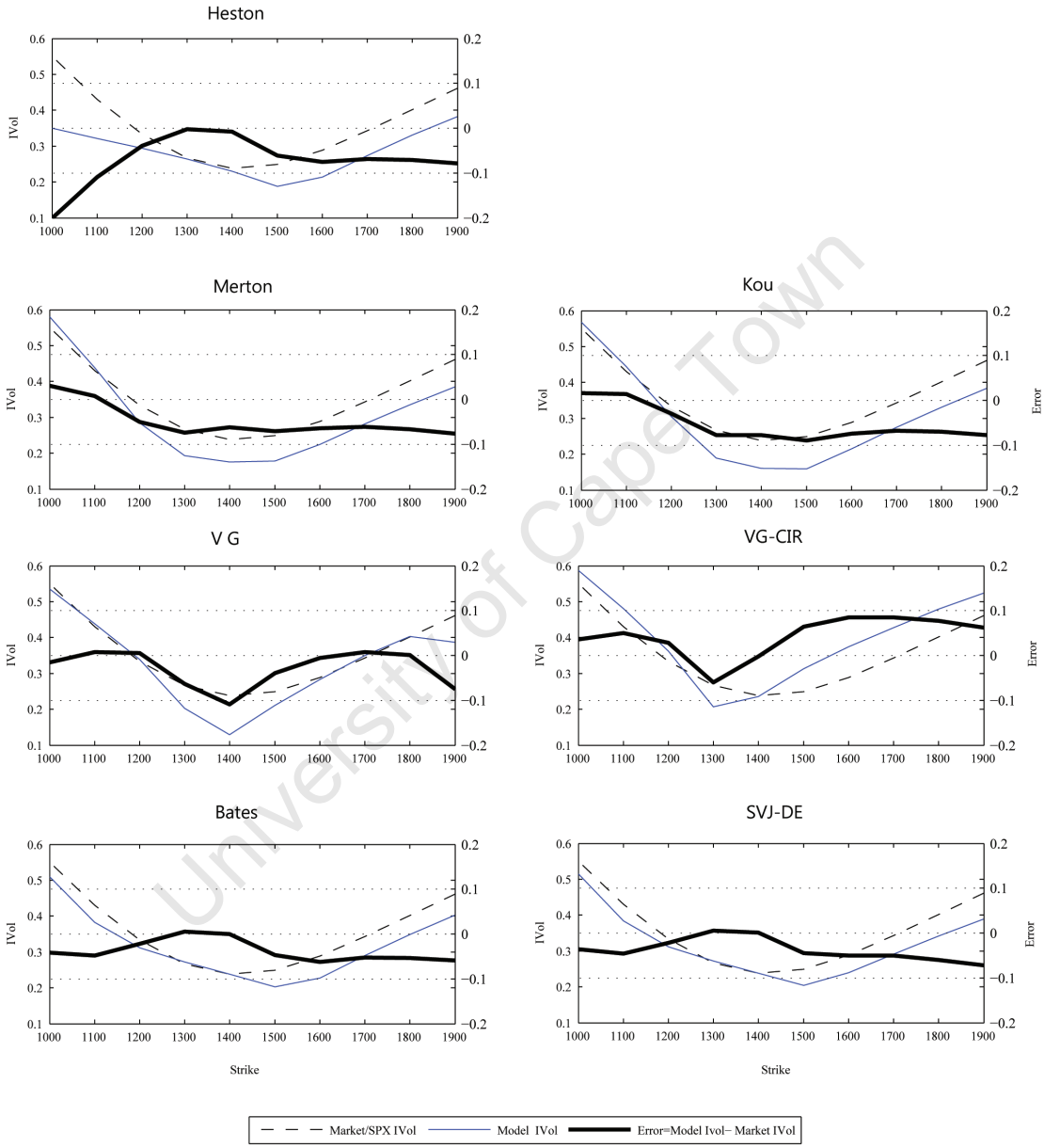



Figure 7.18: Short end implied volatility time slices (T=0.07)

## 7.3 Exotic Option Pricing

Once an option pricing model has been calibrated to the market, it can be used for the pricing of exotic options via Monte Carlo simulation. Barrier call option prices will be computed for the best performing models; that is, the SV and SVJ models. The Heston model will be discretized according to equation 2.27, the Bates model will be discretized according to equation 5.14 and the SVJ-DE model as per equation 5.24. Each simulation was done with 250 time steps and using pseudo-random numbers generated by Matlab. The results were computed using 1 million simulations per option price for the Heston model and $500,000$ simulations per option price for the SVJ models.

If an option has payoff given by $g(S)$, then the Monte Carlo price of an option can be calculated probabilistically via a risk neutral expectation:

$$\begin{aligned} \hat{C} &= e^{-rT}\mathbb{E}_{\mathbb{Q}}[g(S_T)] \\ &= \frac{1}{N}\sum_{i=1}^{N} f(S_i) \end{aligned}$$

where $f(S) = e^{-rT}g(S)$. The sample variance can be calculated

$$s^2(N) = \frac{e^{-2rT}}{N-1}\left(\sum_{i=1}^{N} g^2(S_i) - \frac{1}{N}\left(\sum_{i=1}^{N} g(S_i)\right)^2\right)$$

and then the standard error is given by

$$\frac{s(N)}{\sqrt{N}}.$$



**Barrier Options**

Knock-in options have zero payoff until the asset price crosses a barrier level, then the option has payoff equal to a standard call option. Knock-out options behave like regular options unless the underlying prices crosses the barrier level in which case the option payoff becomes zero. The reader is referred to Schoutens et al. (2003) for a comparison of exotic option prices under Lévy models. The following variants of barrier options are available with associated payoff profiles:

An up-and-out call (UOC) knocks out when it is in the money and hence is also known as a live-out option. It has payoff given by

$$(S_T - K)^+ \mathbf{1}_{\max S_T < B}$$

An up-and-in call (UIC) knocks in when the underlying price crosses a barrier, it has payoff

$$(S_T - K)^+ \mathbf{1}_{\max S_T \geq B}$$

Also note that and UOC plus and UIC has payoff equal to a simple vanilla call option.

An down-and-out call (DOC) knocks out when the asset price drops below a certain level. It

has payoff given by

$$(S_T - K)^+ \mathbf{1}_{\min S_T > B}$$

An down-and-in call (DIC) knocks in when the underlying price drops below a barrier, it has payoff

$$(S_T - K)^+ \mathbf{1}_{\min S_T < B}$$

Similarly, DOC + DIC = vanilla call.

Figure 7.19 shows the prices of up-and-out (live-out) barrier options given by the three best performing models (the SV and SVJ models); figure 7.20 shows the prices of down-and-out barrier options. The barrier level is a multiple of the current spot price. The Heston model did not calibrate to the market data as well as the SVJ models, therefore a difference in exotic option prices is expected. However, the fact that both the SVJ models had similar calibration results (similar fits) does not necessarily imply that they will have similar exotic prices, even thought they do in this case. The SVJ-DE model has the benefit of speed when pricing via Monte Carlo since the algorithm for sampling from a double exponential distribution was found to run faster than the algorithm used to sample from the normal distribution (note that this is a platform and algorithm specific result).

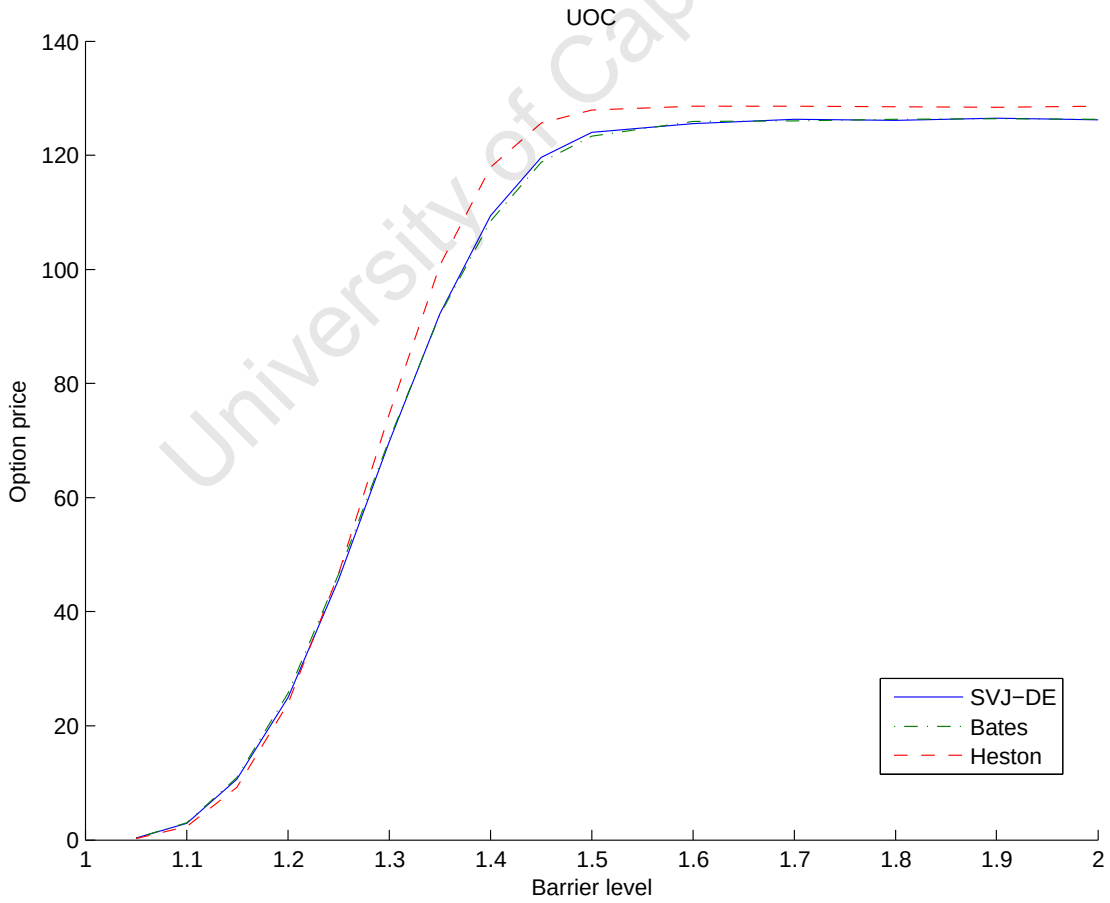



Figure 7.19: Up-and-Out Barrier Prices

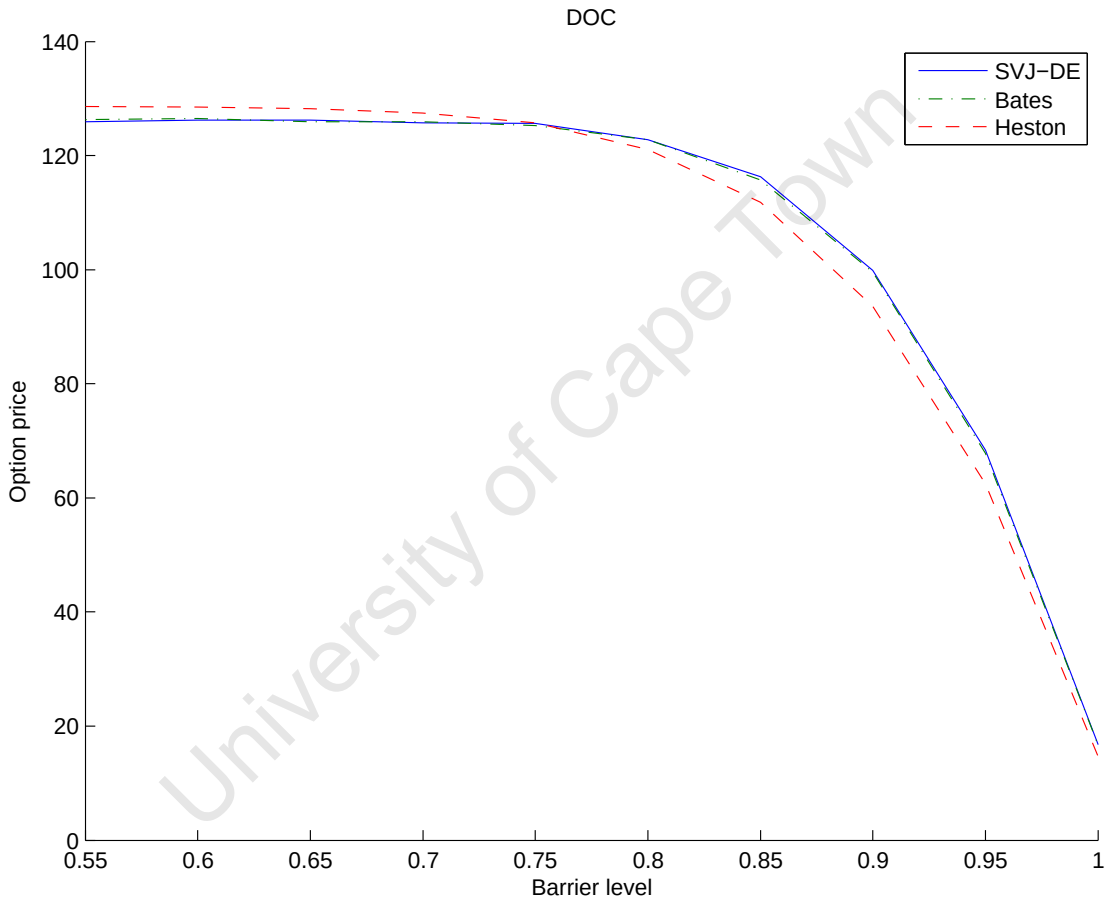





Figure 7.20: Down-and-Out Barrier Prices

# Chapter 8

# Conclusion

This dissertation compared the performance of stochastic volatility models, jump diffusion models, pure jump models and stochastic volatility with jumps (SVJ) models. A new SVJ model was introduced, which combined Heston-type stochastic volatility with Kou-type double exponential jumps, and is abbreviated as the SVJ-DE model. Pricing performance was judged based on the instantaneous fit of the model prices to SPX option prices (model error).

Numerically, the SVJ models were the best performing models; the SVJ-DE model resulted in a smaller error than the Bates model. The Heston SV model gave competitive performance, considering it is defined by only 5 parameters.

The jump-diffusion models and pure jump models gave much larger errors than the SV and SVJ models, resulting in relatively poor performance. The introduction of stochastic time into the VG model in the form of a Cox-Ingersoll-Ross process (to give the VG-CIR model) resulted in an improvement in performance, but the VG-CIR model still performed significantly worse than the SV and SVJ models.

Comparison of market implied volatility surfaces with those generated by the fitted models gave additional insight into model performance. The jump-diffusion models and the pure jumps models showed the ability to generate smile effects, but an inability to generate skew effects. The Heston SV model was able to generate enough skew but not enough smile. The SVJ models were able to adequately capture implied volatility dynamics (smile and skew effects) since they contain both SV and jump components.

Prices of barrier options were obtained from the SV and SVJ models, pricing via Monte Carlo. Both SVJ models gave similar results, with the Heston SV model giving slightly different prices (as expected due to the different fit and different dynamics).

Empirically, underlying asset prices move by jumps (and so are not necessarily continuous); the underlying asset return processes also exhibit stochastic volatility effects. The SVJ models incorporate a jump component as well as a stochastic volatility diffusion component, so that they have intuitively acceptable dynamics; and the models give better numerical performance than the stochastic VG-CIR model. The SVJ models are thus found to give the best overall performance based on model dynamics and numerical performance.

The SVJ-DE model is seen as a superior model (relative to the Bates model) based on numerical

performance, additional flexibility of the jump component, and speed of Monte Carlo simulation.

# Appendix A

# Option Price Data



| Strike Price \ YtM | 0.07 | 0.17 | 0.24 | 0.32 | 0.42 | 0.67 | 0.93 | 1.43 | 1.94 | 2.95 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1000 | | 310.5 | 313.9 | 316.8 | 321 | 330.1 | 338 | 353.5 | 367.3 | 397.5 |
| 1050 | | 263.1 | 268 | 271.9 | 277.2 | 288.5 | 298.1 | | 331.8 | 365.1 |
| 1100 | | 217.2 | 223.7 | 228.8 | 235.1 | 248.6 | 259.8 | 280.1 | 297.7 | 333.8 |
| 1150 | 163.3 | 173.4 | 181.5 | 187.8 | 195.1 | 210.6 | 223.4 | 245.9 | 265.2 | 303.9 |
| 1200 | 119 | 135.8 | 141.9 | 149.3 | 159.5 | 174.9 | 189 | 213.4 | 234.3 | 275.3 |
| 1225 | 98.6 | 116.7 | 123.2 | | 139.7 | 157.9 | 172.7 | 198 | | |
| 1250 | 79.1 | 98.7 | 105.6 | 113.9 | 125 | 141.8 | 157.1 | 183 | 205.2 | 248.1 |
| 1300 | 45 | 62 | 73.5 | 82.4 | 91.7 | 111.7 | 127.8 | 154.8 | 178 | 222.6 |
| 1325 | 32.1 | 47.9 | 59.4 | 68.4 | 77.7 | 97.9 | 114.3 | | | |
| 1350 | 19.75 | 35.7 | 46.9 | 55.7 | 64.9 | 85.1 | 101.5 | 129 | 152.8 | 198.2 |
| 1375 | 11.25 | 25.35 | 35.8 | 44.4 | 53.3 | 73.2 | 89.5 | 117 | | |
| 1400 | 6.2 | 17 | 26.45 | 34.4 | 42.35 | 62.3 | 78.4 | 105.7 | 129.7 | 175.6 |
| 1425 | 2.7 | 10.7 | 18.6 | 25.95 | 33.8 | 52.4 | 68.1 | 95.1 | 119 | 164.8 |
| 1450 | 1.225 | 6.1 | 12.7 | 17.65 | 25.95 | 44.5 | 58.4 | 85 | 108.9 | 154.5 |
| 1475 | 0.6 | 3.325 | 8.2 | 13.4 | 19.6 | 35.6 | 51.3 | | | |
| 1500 | 0.525 | 1.8 | 4.9 | 9.1 | 13.45 | 28.8 | 42.55 | 66.8 | 90.2 | 135.1 |
| 1525 | 0.175 | 1.1 | 3 | | 10.2 | 22.85 | | 58.8 | 81.6 | |
| 1550 | 0.175 | 0.75 | 1.825 | 3.7 | 6.9 | 17.9 | 29.4 | 51.6 | 73.4 | 117.3 |
| 1600 | 0.125 | 0.55 | 0.525 | 1.25 | 2.95 | 10.4 | 19.3 | 38.9 | 59 | 101.1 |
| 1650 | | 0.275 | 0.475 | 0.6 | 1.55 | 5.6 | | 28.3 | | |
| 1700 | | 0.125 | | | 0.65 | 3 | 8.2 | 20.3 | 36.4 | 73.1 |
| 1800 | | | | | 0.475 | 1.075 | 2.5 | 9.5 | 20.95 | 51.5 |
| 1900 | | | | | 0.375 | 0.5 | 1.075 | 3.9 | 11.3 | 34.8 |

Table A.1: Calibration Data: SPX call option price grid (YtM=Years to Maturity)

# Appendix B

# Barrier Option Prices

Up-and-out (live-out) barrier option prices:

| | SV | SVJ | SVJ |
|---|---|---|---|
| Barrier | Heston | Bates | SVJ-DE |
| 1.05 | 0.2250 | 0.2968 | 0.3009 |
| 1.1 | 2.3108 | 2.9939 | 2.8982 |
| 1.15 | 9.1694 | 10.9213 | 10.6514 |
| 1.2 | 23.8876 | 25.6800 | 24.9196 |
| 1.25 | 46.7494 | 46.6519 | 45.6558 |
| 1.3 | 74.7170 | 70.1678 | 69.8469 |
| 1.35 | 100.6969 | 92.0129 | 92.1938 |
| 1.4 | 117.8906 | 108.4163 | 109.3947 |
| 1.45 | 125.5945 | 118.7835 | 119.6142 |
| 1.5 | 127.9325 | 123.2995 | 124.0385 |
| 1.6 | 128.6009 | 125.9037 | 125.5796 |
| 1.7 | 128.5600 | 126.0018 | 126.3014 |
| 1.8 | 128.5328 | 126.2935 | 126.1347 |
| 1.9 | 128.4218 | 126.4383 | 126.4586 |
| 2 | 128.5783 | 126.3353 | 126.1649 |

Down-and-out barrier option prices:

| | SV | SVJ | SVJ |
|---|---|---|---|
| Barrier | Heston | Bates | SVJ-DE |
| 0.55 | 128.5974 | 126.3398 | 125.8786 |
| 0.6 | 128.4884 | 126.4500 | 126.2426 |
| 0.65 | 128.2617 | 125.9234 | 126.1921 |
| 0.7 | 127.4507 | 125.9637 | 125.7774 |
| 0.75 | 125.6853 | 125.2658 | 125.6419 |
| 0.8 | 121.0994 | 122.7437 | 122.7917 |
| 0.85 | 111.7658 | 115.6801 | 116.3144 |
| 0.9 | 93.5493 | 99.5762 | 99.8708 |
| 0.95 | 62.4297 | 67.8169 | 68.3702 |
| 1 | 14.6893 | 16.6285 | 16.7360 |

# Appendix C

# M-Files

Matlab files can be found on a cd accompanying this dissertation. These include input files, function m-files, script m-files and output files. The files are divided into the following folders:

- **0Data** - Contains main input file **marketdata.mat**. Also includes additional m-files for axes labeling, and the SPX volatility surface.
- **0SPXFullSurfPlot** - SPX volatility surface data plus script to generate surface.
- **asamin** - the ASA c-language source plus the asamin interface files.
- **BS** - Black-Scholes call option price function and implied volatility calculator function.
- **D Heston** - SV diffusion model.
- **J VG** - Jump model.
- **JD Kou** - Jump diffusion model.
- **JD Merton** - Jump diffusion model.
- **SVJ Bates** - SVJ jump diffusion model.
- **SVJ DE** - SVJ jump diffusion model.
- **SVJ VGCIR** - jump model with stochastic time.

Each model folder has two subfolders, which may contain some or all of the following subsubfolders:

- **mfiles** - pricing functions and scripts
  - *ASA* - global optimization functions and scripts.
  - *lsqnonlin* - local optimization functions and scripts.
  - *MC_exotics* - Monte Carlo functions and scripts, including barrier option files.
- **output** - calibration output files and exotic option output files.

# Appendix D

# Itô's Formula

Several versions of Itô's formula are presented below and are taken from Hunt and Kennedy (2004).

**Theorem D.1** (Itô's Formula)**.** *Let $f : \mathbb{R} \to \mathbb{R}$ be $C^2$ and let $X$ be a continuous semi-martingale. Then, almost surely, for all $t \geq 0$,*

$$f(X_t) = f(X_0) + \int_0^t f'(X_u)dX_u + \frac{1}{2}\int_0^t f''(X_u)d[X]_u. \tag{D.1}$$

*In particular, if $X$ has the decomposition $X = X_0 + M + A$ then $f(X_t)$ has the decomposition*

$$f(X_t) = f(X_0) + \int_0^t f'(X_u)dM_u + \left[\int_o^t f'(X_u)dA_u + \frac{1}{2}\int_0^t f''(X_u)d[M]_u\right],$$

*and is thus a continuous semi-martingale.*

D.1 can be re-written in differential notation:

$$df(X) = f'(X)dX + \frac{1}{2}f''(X)d[X],$$

**Definition D.2.** A process $X = (X^{(1)}, \ldots, X^{(n)})$ defined relative to $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ with values in $\mathbb{R}^n$ is called a continuous semi-martingale if each coordinate process $X^{(i)}$ is a continuous semi-martingale.

**Theorem D.3** (Multi-dimensional Itô's Formula)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be $C^2(\mathbb{R}^n)$ and let $X = (X^{(1)}, \ldots, X^{(n)})$ be a continuous semi-maringale in $\mathbb{R}^n$. Then, almost surely, $f(X_t)$ is a continuous semi-martingale and*

$$f(X_t) = f(X_0) + \sum_{i=1}^n \int_0^t \frac{\partial f}{\partial x_i}(X_u)dX_u^{(i)} + \frac{1}{2}\sum_{i,j=1}^n \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_u)d[X^{(i)}, X^{(j)}]_u.$$

# Appendix E

# Glossary of Terms and Definitions



**Option** Financial contract which allows the buyer of the option the right to buy (call option) or sell (put option) the underlying asset at an agreed price (the strike price) at an agreed future date (the option expiry date). The writer of the option is therefore obliged to sell or buy the asset, respectively. A type of contingent claim.

**Implied Volatility** Consider a particular option contract, assume that strike price, time to expiry, underlying spot price, risk-free rate and dividend yield are known. Then the implied volatility of an option is the volatility which, when input into the Black-Scholes option pricing model, will give the market price of the option. It is commonly described as "the wrong number, input into the wrong model, to give the right price".

**Put-Call Parity** This is the relationship between the price of a call option and the price of a put option on the same underlying asset; it is given by $Price_{put} + Price_{underlying} = Strike \cdot e^{-rt} + Price_{call}$.

**Moneyness** The relates to the strike price of the option(*K*) and to the spot price of the underlying (*S*). Consider a call option, the call is said to be in the money if $K < S$, at the money if $K = S$ and out of the money if $S < K$.

**Filtration** Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a filtration is a sequence of non-decreasing $\sigma$-algebra's given by $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ with $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$ for $0 \leq s < t < \infty$.

**Adapted Stochastic Process** A stochastic process $X = \{X_t\}_{t\geq 0}$, which is defined on a filtered probability space (or stochastic base) given by $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, is said to be adapted if $X_t$ is $\mathcal{F}_t$ measurable (i.e. $X_t \in \mathcal{F}_t$) for all $t \geq 0$.

**Martingale** (Steele, 2001) A stochastic process $\{X_t : 0 \leq t < \infty\}$ defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$ is said to be a martingale if the following conditions are satisfied:

- $\{X_t\}$ is adapted to $\{\mathcal{F}_t\}$
- $\mathbb{E}\{|X_t|\} < \infty$ for all $0 \leq t < \infty$
- $\mathbb{E}\{X_t|\mathcal{F}_s\} = X_s$ for all $0 \leq s \leq t < \infty$

**Local Martingale** (Steele, 2001, §7.2) If a process $\{M_t\}$ is adapted to the filtration $\{\mathcal{F}_t\}$ for all $0 \leq t < \infty$, then $\{M_t : 0 \leq t < \infty\}$ is called a local martingale provided that there is a non-decreasing sequence $\{\tau_k\}$ of stopping times with the property that $\tau_k \to \infty$ with probability one as $k \to \infty$ and such that for each $k$ the process defined by

$$M_t^{(k)} = M_{t \wedge \tau_k} - M_0 \qquad \text{for } t \in [0, \infty) \tag{E.1}$$

is a martingale with respect to the filtration $\{\mathcal{F}_t : 0 \leq t < \infty\}$.]

**Semi-Martingale** (Hunt and Kennedy, 2004, §3.83) A process $X$ is called a semi-martingale, relative to the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$, if $X$ is an adapted process which can be written in the form

$$X_t = X_0 + M + A \tag{E.2}$$

where $X_0$ is an $\mathcal{F}_0$-measurable random variable, $M$ is a local martingale null at zero and $A$ is an adapted càdlàg process, also null at zero, having paths of finite variation.

**càdlàg , càglàd** A càdlàg process is defined as a right continuous process with left limits (RCLL); whilst a càglàd process is left continuous with right limits.

**Finite variation** If $f : [0, \infty) \to \mathbb{R}^d$ is a function, and given the interval $[a, b]$, then the total variation of $f$ over the interval is defined by

$$V(f; [a, b]) = \sup \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|$$

where the supremum is taken over all finite partitions $0 \leq a = t_0 < t_1 < \cdots < t_n = b < \infty$ of the interval. If the total variation is finite, then the function is said to be of finite variation over the interval.

**Brownian Motion** (Steele, 2001) A continuous time stochastic process $\{B_t : 0 \leq t < T\}$ is called a standard Brownian motion on $[0, T)$ if it has the following four properties:

(i) $B_0 = 0$.

(ii) The increments of $B_t$ are independent; that is, for any finite set of times $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n \leq T$ the random variables

$$B_{t_2} - B_{t_1}, B_{t_3} - B_{t_2}, \cdots, B_{t_n} - B_{t_{n-1}}$$

are independent.

(iii) For any $0 \leq s \leq t < T$ the increment $B_t - B_s$ has the Gaussian distribution with mean 0 and variance $t - s$.

(iv) For all $\omega$ in a set of probability one, $B_t(\omega)$ is a continuous function of $t$ (or $B_t$ has continuous sample paths almost surely).

**Equivalent Measures** Given two probability measures, $\mathbb{P}$ and $\mathbb{Q}$, defined on the measurable space $(\Omega, \mathcal{F})$. Then:

(i) $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$, written $\mathbb{Q} \ll \mathbb{P}$, if for all $F \in \mathcal{F}$

$$\mathbb{P}(F) = 0 \Rightarrow \mathbb{Q}(F) = 0.$$

(iii) $\mathbb{P}$ and $\mathbb{Q}$ are said to be equivalent, written $\mathbb{P} \sim \mathbb{Q}$, if both

$$\mathbb{P} \ll \mathbb{Q} \text{ and } \mathbb{Q} \ll \mathbb{P}.$$

**Numéraire** A numéraire is price process $X_t$ which satisfies $X_t > 0$ almost surely.

**Self-financing** Bingham and Kiesel (2004, §6.1.2)

(i) The value of the portfolio $\varphi$ at time $t$ is given by the scalar product

$$V_\varphi(t) := \varphi(t) \cdot S(t) = \sum_{i=0}^{d} \varphi_i(t) S_i(t), \qquad t \in [0, T]. \tag{E.3}$$

The process $V_\varphi(t)$ is called the *value process*, or wealth process, of the trading strategy $\varphi$.

(ii) The *gains process* $G_\varphi(t)$ is defined by

$$G_\varphi(t) := \int_0^t \varphi(u) dS(u) = \sum_{i=0}^{d} \int_0^t \varphi_i(u) dS_i(u). \tag{E.4}$$

(iii) A trading strategy $\varphi$ is called *self-financing* if the wealth process $V_\varphi(t)$ satisfies

$$V_\varphi(t) = V_\varphi(0) + G_\varphi(t) \qquad \text{for all } t \in [0, T]. \tag{E.5}$$

A self-financing trading strategy requires that the value of the portfolio at time *t* be equal to the initial value of the portfolio plus the gains made up until time *t*; there can be no additions to, or subtractions from, the portfolio.

# Bibliography


Albrecher, H., Mayer, P., Schoutens, W., and Tistaert, J. (2006). The little heston trap. *Working Paper*.

Attari, M. (2004). Option pricing using fourier transforms: A numerically efficient simplification. *Working Paper*.

Bakshi, G. and Madan, D. (2000). Spanning and derivative valuation. *Financial Economics*, (55):205–238.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, 9(1):69–107.

Bingham, N. H. and Kiesel, R. (2004). *Risk Neutral Valuation - the pricing and hedging of financial derivatives*. Springer Finance, London, second edition.

Bjork, T. (2004). *Arbitrage Theory in Continuous Time*. Oxford University Press, second edition.

Broadie, M. and Kaya, O. (2004). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Working Paper*.

Carr, P., Geman, H., Madan, D., and Yor, M. (2001). Stochastic volatility for levy processes. *EFA 2002 Berlin Meetings Presented Paper. Available at SSRN*.

Carr, P., Geman, H., Madan, D., and Yor, M. (2002). The fine structure of asset returns: An empirical investigation. *Journal of Business*, 75(2):305–332.

Carr, P. and Madan, D. (1998). Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2(4):61–73.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236.

Cont, R. and Fonseca, J. (2002). Dynamics of implied volatility surfaces. *Quantitative Finance*, 2(1):45–60.

Cont, R. and Tankov, P. (2003). *Financial modelling with Jump Processes*. Chapman and Hall Crc Financial Mathematics Series.

Cox, J., Ingersoll, J., and Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica*, (53):385–408.

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 2(1994):61–73.

Delbaen, F. and Schachermayer, W. (1998). The fundamental theorem of asset pricing for unbounded stochastic processes. *Mathematische Annalen*, 312(1998):215–250.

Derman, E. and Kani, I. (1998). Riding on a smile. *Risk*, 7(2).

Derman, E., Kani, I., and Chriss, N. (1996). Implied trinomial trees of the volatility smile. *Journal of Derivatives*, 4.

Dumas, B., Fleming, J., and Whaley, R. (1998). Implied volatility functions: Empirical tests. *The Journal of Finance*, LIII(6):2059–2105.

Dupire, B. (1994). Pricing with a smile. *Risk*, 7(1).

Gatheral, J. (2004). A parsimonious arbitrage-free implied volatility parameterization with application to the valuation of volatility derivatives. www.math.nyu.edu/fellows-fin-math/gatheral/madrid2004.pdf.

Gatheral, J. (2006). *The Volatility Surface: A Practitioners Guide*. John Wiley and Sons, New Jersey, first edition.

Hagan, S., Kumar, D., Lesniewski, A., and Woodward, D. (2002). Managing smile risk. *Wilmott Magazine*, 2002(September):84–108.

Hamida, S. and Cont, R. (2005). Recovering volatility from option prices by evolutionary optimization. *Journal of Computational Finance*, 8(4).

Heston, S. L. (1993). A closed form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343.

Hull, J. and White, A. (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, 42(2):281–300.

Hunt, P. and Kennedy, J. (2004). *Financial Derivatives in Theory and Practice*. John Wiley and Sons, England, revised edition.

Kahl, C. and Jackel, P. (2006). Not-so-complex logarithms in the heston model. *Working Paper*.

Kilin, F. (2007). Accelerating the calibration of stochastic volatility models. *Working Paper*, Munich Personal RePEc Archive.

Kou, S. G. (2002). A Jump-Diffusion Model for Option Pricing. *Management Science*, 48(8):1086–1101.

Kou, S. G. and Wang, H. (2004). Option pricing under a double exponential jump diffusion model. *Management Science*, 50(9):1178–1192.

Lee, R. (1999). Implied and local volatilities under stochastic volatility. *International Journal of Theoretical and Applied Finance*, 4(1):45–89.

Lewis, A. (2001). A simple option formula for general jump-diffusion and other exponential levy processes. *Working Paper*.

Lord, R., Koekkoek, R., and van Dijk, D. (2008). A comparison of biased simulation schemes for stochastic volatility models. *Tinbergen Institute Discussion Paper*, 2006-046/4.

Madan, D., Carr, P., and Chang, E. (1998). The variance gamma process and option pricing. *European Finance Review*, 2:79–105.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144.

Mikhailov, S. and Nogel, U. (2003). Hestons stochastic volatility model implementation, calibration and some extensions. *Wilmott Magazine*, July:74–79.

Moins, S. (2002). Implementation of a simulated annealing algorithm for matlab. *Training Report*.

Moodley, N. (2005). The heston model: A practical approach. *Honours Project*, University of the Witwatersrand, Johannesburg, South Africa.

Sato, K. (1999). *Levy Processes and Infinitely Divisible Distributions*. Cambridge University Press.

Schoutens, W. (2003). *Levy Processes in Finance - Pricing Financial Derivatives*. Wiley, England.

Schoutens, W., Simons, E., and Tistaert, J. (2003). A perfect calibration! now what? *Working Paper*.

Shreve, S. (2004). *Stochastic Calculus for Finance II: Continuous Time Models*. Springer, New York.

Steele, M. J. (2001). *Stochastic Calculus and Financial Applications*. Springer, New York.

Vogt, A. (2005). Gatheralsmile-vola excelonly.zip. http://www.axelvogt.de/axalom/.

# Index