A study of the Hobson and Rogers volatility model

Gearóid Ryan
104521170

NATIONAL UNIVERSITY OF IRELAND, CORK

School of Mathematical Sciences

Department of Mathematics

Thesis submitted for the degree of Doctor of Philosophy

February 2014

Supervisors: Dr. Tom Carroll

Head of Department/School: Prof. Bernard Hanzon

Research supported by Edgeworth Centre for Financial Mathematics

Contents

I, Gearóid Ryan, certify that this thesis is my own work and I have not obtained a degree in this university or elsewhere on the basis of the work submitted in this thesis.

Gearóid Ryan

List of Figures

List of Tables

Chapter 1

Introduction and Literature Review

The aim of this thesis is to investigate complete stochastic volatility models for financial asset prices quoted on the stock market. We are especially interested in the Hobson and Rogers model [1]. A simplified version of the model has some very attractive analytical features which will be discussed. The pricing of options contracts on assets is one of the main uses of these models. In fact, market option price data may be used to calibrate the parameters of such a model. We initially look at the famous option pricing model of Black and Scholes. Techniques used in its derivation are fundamental to modern option pricing methods and we will show parallels between the work of Black and Scholes and that of Hobson and Rogers. Modeling volatility is very important in asset pricing theory. Early theories have presumed the volatility of the asset to be constant. Empirical data shows this not to be the case. Nevertheless, these early theories have formed the basis for more modern work. We will discuss discrete and stochastic volatility models as well as their respective implications for option pricing.

1.1 The Black-Scholes Model

In 1973, Fischer Black and Myron Scholes published 'The Pricing of Options and Corporate Liabilities' [2]. Here, under a number of assumptions about the dynamics of asset prices, a deterministic formula for option prices is derived. This model will be the starting point of this thesis. We look at the underlying assumptions, including that of constant volatility, and provide a derivation of the partial differential equation that describes the option price. This derivation and its assumptions have, however, been the source of some controversy. We present and discuss the original derivation. We solve the equation analytically and discuss approaches taken by Carr, by Beck, by Björk and by Øksendal.

1.1.1 Assumptions and Standard Derivation

We assume that we have a market consisting of an asset or stock whose value is denoted by $X(t)$, a risk free asset, or bond, whose value is denoted by $B(t)$ and a European call option, or asset derivative, whose value is assumed to depend on the asset value and time to expiry. The most basic assumption is that our asset value follows Geometric Brownian Motion.

$$dX(t) = \mu X(t)dt + \sigma X(t)dW(t) \tag{1.1}$$

where μ and σ are constants and $W(t)$ is a Brownian motion. This model implies that the asset returns are log-normally distributed with mean $(\mu - \frac{\sigma^2}{2})t$ and variance $\sigma^2 t$. Financial terminology refers to σ as the 'volatility'. It is well known that this model is not accurate in that the implication of log-normally distributed returns does not hold in reality. Historical market data has shown that there is a greater probability of very positive or very negative values than predicted by the model. This empirical data is said to have 'fat tails', a

reference to the shape of the distribution curve of asset returns. Black and Scholes describe how to construct a portfolio consisting of assets and options that replicates the value of a bond at expiry. A no-arbitrage argument then shows that the portfolio value must always equal the value of the bond at the time of writing. This allows us to derive a partial differential equation for the value of the option which, in this case, can be solved explicitly giving the value of the option. The argument is motivated by a discrete model of how the value of a replicating portfolio changes in time.

1.1.2 The standard portfolio derivation

Construct a portfolio by selling a call option and holding $Q(t)$ risky assets.[1]

Definition 1 (European Call Option)**.** *A European call option is a contract where the holder of the option has the right but not the obligation to buy the underlying asset $X(t)$ at a specified future date T for a specified future (strike) price K. At maturity the owner of the option will have to decide whether or not to exercise the call option.*

At maturity, the value of the option will depend on the value of the underlying asset relative to the strike price.

1. $X(T) < K$: In this case the option is valueless as the owner of the option will not want to buy the asset for K if it can be bought on the market more cheaply for $X(T)$. The option will not be exercised and is worth zero.

2. $X(T) > K$: The option holder will exercise the option, buying the asset at price K, then potentially immediately selling the asset on the market for price $X(T)$. In this case the option is worth $X(T) - K$.

[1] Standard notation such as that found in [3] uses $\Delta(t)$ to denote the number of assets held in the portfolio. Since we prefer to reserve this label for time intervals we will use $Q(t)$ instead.

The value of an European call option at maturity is therefore given by

$$\max\{X(T) - K, 0\} \text{ or } (X(T) - K)^+.$$

Some implicit assumptions are made:

1. The derivative instrument is question can be bought and sold on a market.
2. The market is free of arbitrage.
3. The price process for the derivative asset, i.e. the option, is of the form

$$f\Big(t, X(t)\Big)$$

where f is some smooth function.

The main idea is that the relative holdings of the option and asset in the portfolio are varied with time in such a way that the overall portfolio value is risk free and thus grows at the same rate as a bond. In other words there is some relative weighting of positions in the asset and option such that random fluctuations in the value of each will cancel each other out. Maintaining this weighting is known as hedging. If a perfect hedge is created then the portfolio is completely risk free as there is no random element.

To see that this is possible consider the following: If the value of an asset increases then so does a call option written on that asset. We can justify this statement by saying that if the asset value is higher then it is more likely that $X(T) > K$. If we buy an asset and sell an option it is clear that any increase in the value of our asset position value will correspond to a decrease in value of our option position and visa versa. Assume now that we know the relative weighting $Q(t)$ to maintain the risk free status. We have assumed that there are no arbitrage opportunities in our market. This means that the value of a

risk-free portfolio from which funds are not withdrawn or added (self-financing) must grow at the same rate as a risk-free bond. The value process of the portfolio $V(t)$ is given by

$$V(t) = -f\Big(t, X(t)\Big) + QX(t). \tag{1.2}$$

The work of Black and Scholes tells us this portfolio is risk free and self-financing and thus replicates a bond by the choice[2] $Q = \frac{\partial f(t,X(t))}{\partial X}$. In order to check if the portfolio is risk free we compute the differential of the portfolio using Itô's formula. We are required to do this since the value of the option and asset positions in the portfolio are random variables. We find that

$$\begin{aligned} dV\Big(t, X(t)\Big) &= -f_t dt - \frac{1}{2} f_{XX}\Big(dX(t)\Big)^2 - f_X dX(t) + Q dX(t) \\ &= -\Big(f_t + \frac{1}{2}\sigma^2 X^2 f_{XX}\Big)dt, \end{aligned} \tag{1.3}$$

using (1.1) and $\Big(dW(t)\Big)^2 = t$. We see that the above choice of Q had ensured that the stochastic components of the portfolio have been eliminated. Since the portfolio is deterministic and we have assumed there are no arbitrage opportunities in our market, our portfolio must grow at the same rate as a bond. The dynamics of a bond $B(t)$ are given by

$$\begin{aligned} dB(t) &= rB(t)dt \\ B(t) &= e^{-r(T-t)} \end{aligned}$$

where we are assuming for now that the risk-free interest rate is constant and that the bond takes the value 1 at some future time $T \geq t$. The return on our

[2] In reality, due to the discrete nature of trading, this can never be the case.

risk free portfolio is

$$\begin{aligned} dV\Big(t, X(t)\Big) &= rV\Big(t, X(t)\Big)dt. \\ \Rightarrow -\Big(f_t + \frac{1}{2}\sigma^2 X^2 f_{XX}\Big)dt &= r\Big(- f(t, X(t)) + f_X X(t)\Big)dt \\ \Rightarrow f_t + rXf_X + \frac{1}{2}\sigma^2 X^2 f_{XX} - rf &= 0, \end{aligned}$$

which is the Black-Scholes equation. We can solve this for f giving a formula for the option price.

1.1.3 Problems with the standard derivation

The above derivation leaves a number of questions unanswered.

1. Why don't we treat Q as a stochastic variable when computing $dV(t, X(t))$?
2. In reality, assets and options are traded in discrete time, while the above argument assumes continuous trading. What effect will this have on the derivation?
3. We required that in order for $dV\Big(t, X(t)\Big) = rV\Big(t, X(t)\Big)dt$ that the portfolio be self-financing. How do we define self-financing mathematically and does the Black-Scholes portfolio meet the self-financing requirement?

We will now present a more rigorous derivation and deal will all of the above points simultaneously. We will consider a portfolio consisting of $\alpha(t, X(t))$ options, $Q(t, X(t))$ risky assets and $\beta(t, X(t))$ bonds. We only consider strategies that are 'admissible' as defined below. We will treat α, Q and β as random variables. We will use a discrete argument to motivate a definition of self-financing and discuss the self-financing requirement. We will conclude that the Black-Scholes portfolio described above is not self-financing, invalidating the

traditional derivation. We will discuss other approaches in the literature, and provide necessary conditions for all risk free self-financing portfolios.

Definition 2 (Admissible Strategy)**.** *A strategy (α, β) is admissible if it is bounded from below, i.e. there exists a constant C such that*

$$V_t^{(\alpha,\beta)} \geq C,$$

for all $t \geq 0$.

1.1.4 Self-financing condition

We now consider more carefully the concept of self-financing in the context of the traditional portfolio consisting of a stock, an option on that stock and a risk free bond, as considered in Section 1.1.2.

Definition 3 (Itô Process)**.** *An Itô Process is defined to be an adapted stochastic process that can be expressed as the sum of an integral with respect to Brownian motion and an integral with respect to time.*

$$X_t = X_0 + \int_0^t \sigma_s dB_s + \int_0^t \mu_s ds.$$

Here, B is a Brownian motion and it is required that σ is a predictable $B-$integral process, and μ is predictable and Lebesgue integrable.

Treating all our variables as Itô process's, we recall the product rule for stochastic variables [4], which follows from Itô's formula, in the computation at equation (1.3).

Lemma 1 (Product rule for Itô processes[3])**.** *Given two Itô process's*

[3]We use notation from [5], Chapter 4.

$(X(t), Y(t))$, *let* $g(X(t), Y(t)) = X(t)Y(t)$. *Now using Itô's formula we obtain*

$$\begin{aligned} d(g(X(t), Y(t))) &= \frac{\partial g}{\partial X}dX + \frac{\partial g}{\partial Y}dY + \frac{1}{2}\frac{\partial^2 g}{\partial X^2}d[X, X] \\ &\quad + \frac{\partial^2 g}{\partial X \partial Y}d[X, Y] + \frac{1}{2}\frac{\partial^2 g}{\partial Y^2}d[Y, Y] \\ &= XdY + YdX + d[X, Y]. \end{aligned}$$

Proposition 2 (Self-financing condition)**.** *Consider a portfolio consisting of* α *options,* Q *stocks and* β *bonds. The value of our portfolio is given by*

$$V\Big(t, X(t)\Big) = \alpha\Big(t, X(t)\Big)f\Big(t, X(t)\Big) + Q\Big(t, X(t)\Big)X(t) + \beta\Big(t, X(t)\Big)B(t)$$

Such a portfolio is self-financing if

$$d\alpha\,(f + df) + dQ\,(X + dX) + d\beta\,(B + dB) = 0. \tag{1.4}$$

Proof. Firstly we consider what is physically happening when we hedge our portfolio. We will drop the arguments from here on. Preceding each time period $[t, t + \Delta_t]$, the values of α, Q, β (called the *trading strategy*) are calculated such that the portfolio is risk free. During each time period $[t, t + \Delta_t]$ the trading strategy is fixed while the values of the option, asset and bond may change. Noting the new values of f, X and B, a new trading strategy is constructed in order to maintain a risk free portfolio. The value of the portfolio at time $t + \Delta_t$ and before hedging is given by

$$V_-(t + \Delta_t) = \alpha(t)f(t + \Delta_t) + Q(t)X(t + \Delta_t) + \beta(t)B(t + \Delta_t).$$

After hedging, the portfolio has the value

$$V_+(t + \Delta_t) = \alpha(t + \Delta_t)f(t + \Delta_t) + Q(t + \Delta_t)X(t + \Delta_t) + \beta(t + \Delta_t)B(t + \Delta_t)$$

If no external funds are to be withdrawn or put into the portfolio during the hedging process then the above expressions need to be equal. This leads to the condition

$$\begin{aligned} f(t+\Delta_t)\Big(\alpha(t+\Delta_t)-\alpha(t)\Big)+X(t+\Delta_t)\Big(Q(t+\Delta_t)-Q(t)\Big) & \\ +B(t+\Delta_t)\Big(\beta(t+\Delta_t)-\beta(t)\Big) &= 0. \end{aligned}$$

Note that in the above expression we have backward differences. Itô calculus requires the use of forward differences. We can achieve this by adding and subtracting the quantity

$$f(t)\Big(\alpha(t+\Delta_t)-\alpha(t)\Big)+X(t)\Big(Q(t+\Delta_t)-Q(t)\Big)+B(t)\Big(\beta(t+\Delta_t)-\beta(t)\Big)$$

to the previous expression. Our final condition is

$$\begin{aligned} &\Big(\alpha(t+\Delta_t)-\alpha(t)\Big)f(t)+\Big(\alpha(t+\Delta_t)-\alpha(t)\Big)\Big(f(t+\Delta_t)-f(t)\Big) \\ &+\Big(Q(t+\Delta_t)-Q(t)\Big)X(t)+\Big(Q(t+\Delta_t)-Q(t)\Big)\Big(X(t+\Delta_t)-X(t)\Big) \\ &+\Big(\beta(t+\Delta_t)-\beta(t)\Big)B(t)+\Big(\beta(t+\Delta_t)-\beta(t)\Big)\Big(B(t+\Delta_t)-B(t)\Big) = 0. \end{aligned}$$

This is the discrete form of the self-financing condition. Note that the above equation uses asset/option/bond values at some time t and the changes in quantities of each held only between t and the end of the next time period i.e. we now only have forward differences. If we take the limit at $\Delta_t \to 0$ we have

$$d\alpha(t)f(t)+d\alpha(t)df(t)+dQ(t)X(t)+dQ(t)dX(t)+d\beta(t)B(t)+d\beta(t)dB(t)=0.$$

Collecting terms, and dropping the arguments for simplicity, gives us an

expression for the continuous self-financing condition, namely

$$d\alpha\,(f + df) + dQ\,(X + dX) + d\beta\,(B + dB) = 0.$$

Therefore, by the product rule, we have the more traditional form,

$$\alpha df + QdX + \beta dB = 0. \qquad \square$$

1.1.5 Risk free condition

Proposition 3. *[Risk free condition for general portfolio] Consider a portfolio consisting of α options, Q stocks, and β bonds. The value of the portfolio is given by*

$$V\Big(t, X(t)\Big) = \alpha\Big(t, X(t)\Big) f\Big(t, X(t)\Big) + Q\Big(t, X(t)\Big) X(t) + \beta\Big(t, X(t)\Big) B(t).$$

Such a portfolio is risk free if

$$\Big(\alpha_X f + Q_X X + \beta_X B\Big) + Q + \alpha f_X = 0. \tag{1.5}$$

If such a portfolio is self-financing then this condition reduces to

$$Q + \alpha f_X = 0. \tag{1.6}$$

Proof. Looking at the differential of the value of the portfolio we have

$$\begin{aligned} dV &= \alpha df + QdX + \beta dB \\ &\quad + d\alpha(f + df) + dQ(X + dX) + d\beta(B + dB). \end{aligned}$$

Using Itô's formula we have

$$\begin{aligned}
d\alpha &= \alpha_t dt + \alpha_X dX + \frac{1}{2}\alpha_{XX}(dX)^2 \\
&= \Big(\alpha_t + rX\alpha_X + \frac{1}{2}\sigma^2 X^2 \alpha_{XX}\Big)dt + \sigma X \alpha_X dW, \\
d\beta &= \beta_t dt + \beta_X dX + \frac{1}{2}\beta_{XX}(dX)^2 \\
&= \Big(\beta_t + rX\beta_X + \frac{1}{2}\sigma^2 X^2 \beta_{XX}\Big)dt + \sigma X \beta_X dW, \\
dQ &= Q_t dt + Q_X dX + \frac{1}{2}Q_{XX}(dX)^2 \\
&= \Big(Q_t + rXQ_X + \frac{1}{2}\sigma^2 X^2 Q_{XX}\Big)dt + \sigma X Q_X dW \\
df &= f_t dt + f_X dX + \frac{1}{2}f_{XX}(dX)^2 \\
&= \Big(f_t + rXf_X + \frac{1}{2}\sigma^2 X^2 f_{XX}\Big)dt + \sigma X f_X dW \\
dX &= rXdt + \sigma X dW, \\
dB &= rBdt.
\end{aligned}$$

In order for our portfolio to be risk free, the coefficients of dW in dV should sum to zero. This gives

$$\begin{aligned}
X\alpha f_X + X\alpha_X f + QX + X^2 Q_X + X\beta_X B &= 0 \\
\Rightarrow Q_X X^2 + \Big(\alpha f_X + \alpha_X f + Q + \beta_X B\Big)X &= 0 \\
\Rightarrow \Big(\alpha_X f + Q_X X + \beta_X B\Big) + Q + \alpha f_X &= 0 \qquad (1.7)
\end{aligned}$$

This is the risk free condition for a portfolio consisting of a bond, a stock and an option. If this portfolio is self-financing, then using (1.4) we have

$$dV = \alpha df + QdX + \beta dB.$$

Now again if we set the sum of coefficients of dW in dV equal to zero we have

$$Q + \alpha f_X = 0.$$

The first three terms on the left hand side of equation (1.7) have disappeared which is equation (1.6). □

With risk free and self-financing conditions clearly defined above we return to our construction of a risk free, self-financing portfolio of options, assets and bonds. For this type of portfolio to be self-financing we need

$$d\alpha(f + df) + dQ(X + dX) + d\beta(B + dB) = 0 \tag{1.8}$$

and, furthermore, for it to be risk free we need

$$\alpha f_X + Q = 0. \tag{1.9}$$

Proposition 4. *A portfolio consisting of α options, Q stocks and β bonds is risk free and self-financing if and only if the triple α, Q, β is chosen such that*

$$Q = -\alpha f_X \tag{1.10}$$

$$\beta = -\frac{1}{B}\alpha(f - f_X X) + \frac{k}{B} \tag{1.11}$$

and

$$\alpha\Big(- Xrf_X - \frac{1}{2}X\sigma^2 f_{XX} - f_t + rf\Big) - rk + k_t = 0, \tag{1.12}$$

where $k = k(t)$ is any function that doesn't depend on X.

Proof. Equation (1.10) is the risk free condition for a self-financing portfolio as derived above. The second and third are found by solving (1.8). We require the stochastic and deterministic parts of this equation to be simultaneously equal to

zero: that the deterministic part vanishes gives

$$\begin{aligned} f\Big(\frac{1}{2}X^2\sigma^2\alpha_{XX} + Xr\alpha_X + \alpha_t\Big) + X^2\sigma^2 f_X\alpha_X \\ +X\Big(\frac{1}{2}X^2\sigma^2 Q_{XX} + XrQ_X + Q_t\Big) + X^2\sigma^2 Q_X \\ +B(\frac{1}{2}X^2\sigma^2\beta_{XX} + Xr\beta_X + \beta_t) &= 0 \end{aligned} \tag{1.13}$$

while the vanishing of the stochastic part gives

$$f\alpha_X + XQ_X + B\beta_X = 0. \tag{1.14}$$

We now have three constraints (1.9), (1.13), (1.14) on the triple (α, Q, β) and we can attempt to solve these equations. Starting with (1.14) we make the substitution $m = \alpha(f - Xf_X)$, such that

$$\begin{aligned} m_X &= \alpha_X(f - Xf_X) + \alpha(-Xf_{XX}) \\ &= f\alpha_X - X\alpha f_{XX} - X\alpha_X f_X \\ &= f\alpha_X + XQ_X. \end{aligned}$$

by (1.9). Then (1.14) becomes

$$m_X = -B\beta_X = -(B\beta)_X. \tag{1.15}$$

Integrating both sides we have

$$m = \alpha(f - Xf_X) = -B\beta + k(t).$$

The above equation tells us that the condition

$$\beta = -\frac{1}{B}\alpha(f - Xf_X) + \frac{k(t)}{B} \tag{1.16}$$

which is equation (2) as stated in the above proposition, is necessary for self-financing. If we make this choice of β, then according to (1.9) our portfolio has the value

$$V = \alpha f + (-\alpha f_X)X - \frac{1}{B}\alpha(f - Xf_X)B + k(t) \equiv k(t). \tag{1.17}$$

With an expression for β determined, we substitute (1.14) into (1.13) to obtain

$$\begin{aligned} f\left(\frac{1}{2}X^2\sigma^2\alpha_{XX} + \alpha_t\right) + X^2\sigma^2 f_X\alpha_X \\ +X\left(\frac{1}{2}X^2\sigma^2 Q_{XX} + Q_t\right) + X^2\sigma^2 Q_X \\ +B(\frac{1}{2}X^2\sigma^2\beta_{XX} + \beta_t) &= 0 \end{aligned} \tag{1.18}$$

and secondly use (1.9) to obtain

$$\begin{aligned} f\left(\frac{1}{2}X^2\sigma^2\alpha_{XX} + \alpha_t\right) + X^2\sigma^2 f_X\alpha_X \\ +X\left(\frac{1}{2}X^2\sigma^2(-\alpha f_X)_{XX} + (-\alpha f_X)_t\right) + X^2\sigma^2(-\alpha f_X)_X \\ +B(\frac{1}{2}X^2\sigma^2\beta_{XX} + \beta_t) &= 0. \end{aligned} \tag{1.19}$$

Finally, from (1.16) we obtain

$$\begin{aligned} \beta_t &= -e^{r(T-t)}\left(f - Xf_X\right)\alpha_t + re^{r(T-t)}\left(f - Xf_X\right)\alpha - e^{r(T-t)}\left(f_t - Xf_{t,X}\right)\alpha \\ &\quad -e^{r(T-t)}(rk - k_t), \end{aligned}$$

and

$$\beta_{XX} = 2Xe^{r(T-t)}f_{XX}\alpha_X - e^{r(T-t)}\left(f - Xf_X\right)\alpha_{XX} + e^{r(T-t)}f_{XX}\alpha + Xe^{r(T-t)}f_{XXX}\alpha,$$

so that equation (1.19) becomes

$$\alpha\left(-Xrf_X - \frac{1}{2}X^2\sigma^2 f_{XX} - f_t + rf\right) - rk + k_t = 0,$$

which is equation (1.12) in the above proposition, concluding the proof. □

All self-financing risk free portfolios must fall into the above regime and their value must be deterministic since for any choice of α we have

$$V = \alpha f - \alpha f_X X - \frac{\alpha}{B}(f - Xf_X)B + \frac{k(t)}{B}B \equiv k(t) \tag{1.20}$$

We are now ready to discuss in detail the traditional portfolio derivation of the Black-Scholes equation.

1.1.6 Standard Portfolio Derivation Revisited

In the standard derivation of the Black-Scholes equation, as presented previously, the portfolio $V(t, X)$ is chosen that $\alpha = -1$, $Q = f_X$ and $\beta = 0$. If we assume that the portfolio is self-financing then, since (1.9) holds, we know this portfolio is risk free. The self-financing assumption is claimed/assumed in most textbooks without justification. In this portfolio the choice of $\beta = 0$ implies, by (1.10), that for self-financing we must have

$$k(t) = -f + Xf_X = V(t) \tag{1.21}$$

and, by (1.12), with $\alpha = -1$,

$$Xrf_X + \frac{1}{2}X^2\sigma^2 f_{XX} + f_t - rf - rk + k_t = 0. \tag{1.22}$$

The first equation tells us that our portfolio must be deterministic. We have no dependence on X. Substituting the first equation into the second we find

$$\begin{aligned} -rk + k_t &= rf - rXf_x - f_t + (Xf_x)_t \\ &= rf - rXf_x - f_t \end{aligned}$$

which when substituted into 1.22 we have $f_{XX} = 0$ which tells us that the value of the option in our portfolio can be at most a linear function of the underlying at all times. This contradicts the standard derivations where it is found that $f_{XX} \neq 0$. We may actually show that the above portfolio is not self-financing. Take the original self-financing statement

$$d\alpha(f + df) + dQ(X + dX) + d\beta(B + dB) = 0$$

and set $d\alpha = d\beta = 0$ since they are both constants. Then we see we must have $dQ = 0$. But since $Q = f_X$ this cannot be true since $dQ \neq 0$ in general. Finally, if we only require the portfolio to be risk free (and not necessarily self-financing), equation (1.5) tells us that

$$\begin{aligned} \Big(Q_X X\Big) + Q - f_X &= 0, \\ \Rightarrow (QX)_X &= f_X, \\ \Rightarrow Q &= \frac{f}{X} \neq f_X. \end{aligned}$$

The traditional Black-Scholes portfolio is not a risk free, self-financing portfolio. Questions that arise are

1. If the traditional portfolio is neither self-financing nor risk free - why are we able to successfully derive the Black-Scholes equation?

   Given that in the traditional derivation the product rule for Itô processes

(See Lemma 1 above) is not implemented correctly, the derivation implicitly assumes self-financing. Once this assumption is made about a portfolio, the Black-Scholes pde is derived. This consequence is stated more formally in the following proposition:

Proposition 5. *The following two conditions are equivalent:*

(a) A portfolio consisting of stocks and bonds is self-financing.

(b) Denote by $f(t, S_t)$ the value of such a portfolio. Then $f(t, S_t)$ satisfies the Black-Scholes p.d.e.

a)⇒b). Denote the position in the stock by α and that in the bond by β. By a) we have that

$$\begin{aligned} df(t, S_t) &= \alpha dS + \beta dB \\ &= (\alpha\mu S + r\beta B)dt + \alpha\sigma S dW. \end{aligned}$$

Using Itô's lemma for the value process of the portfolio we independently have

$$\begin{aligned} df &= f_t dt + f_S dS + \frac{1}{2}\sigma^2 S^2 f_{SS} dt \\ &= (f_t + \mu S f_S + \frac{1}{2}\sigma^2 S^2 f_{SS})dt \\ &\quad + \sigma S f_S dW. \end{aligned}$$

By uniqueness of representation of an Itô process ([6] Proposition 5.3), we have from the dB coefficients that

$$\sigma S f_S = \alpha\sigma S$$

and using this result, comparing the dt coefficients we have

$$f_t + \frac{1}{2}\sigma^2 S^2 f_{SS} = r\beta B \tag{1.23}$$

Rewriting our portfolio in terms of the above expression gives

$$rf = rf_S S + f_t + \frac{1}{2}\sigma^2 S^2 f_{SS} \tag{1.24}$$

which is the Black-Scholes p.d.e.

b)$\rightarrow$ a) Again using Itô's formula we have

$$\begin{aligned} df &= f_t dt + f_S dS + \frac{1}{2}\sigma^2 S^2 f_{SS} dt \\ &= (f_t + \frac{1}{2}\sigma^2 S^2 f_{SS})dt + f_S dS. \end{aligned}$$

Since we are assuming the Black-Scholes p.d.e. holds we can write

$$\begin{aligned} df &= (rf - rf_S)dt + f_S dS \\ &= r\beta dt + \alpha dS \\ &= \beta dB + \alpha dS \end{aligned}$$

which is the condition for self financing. □

2. How do we create a risk free, self-financing portfolio consisting of stocks and options to replicate a bond?

Proposition 6. *The necessary weights, α and Q, needed to form a risk free, self-financing portfolio of options and underlying stocks must be of the form*

$$(\alpha, Q) = \Big(\frac{B(t)}{f - Xf_X}, -\frac{f_X B(t)}{f - Xf_X}\Big), \tag{1.25}$$

where the price process of the portfolio is given by $V(t) = \alpha f + QX$ and

where, as usual, f denotes the value of the option, and X denotes the prices process of the underlying stock. This hedging strategy allows a derivation of the Black Scholes equation.

Proof. Firstly we know that any portfolio that replicates a bond is automatically risk free and self-financing. The portfolio should contain a number of options and stocks, but zero bonds. We'll denote the positions in each of these assets as usual by (α, Q, β). In order to have $\beta = 0$ we should have $k(t) = \alpha(f - Xf_X)$ as required by equation (). Then we have $\alpha = \frac{k(t)}{f - Xf_X}$ and $Q = \frac{-f_X k(t)}{f - Xf_X}$. Thus our portfolio should have value

$$\begin{aligned} V(t) &= \frac{k(t)}{f - Xf_X} f - \frac{f_X k(t)}{f - Xf_X} X \\ &= k(t) \left(\frac{f - Xf_X}{f - Xf_X} \right) \end{aligned}$$

which means we should choose $k(t) = B(t)$ since we want to replicate a bond. Finally our portfolio is given by

$$(\alpha, Q, \beta) = \left(\frac{B(t)}{f - Xf_X}, -\frac{f_X B(t)}{f - Xf_X}, 0 \right).$$

Now, since we know that the portfolio is self-financing, we know that equation (1.12) from Proposition 4 is satisfied. Therefore we have

$$\begin{aligned} \alpha \left(-Xrf_X - \frac{1}{2} X^2 \sigma^2 f_{XX} - f_t + rf \right) - rk + k_t &= 0 \\ \Rightarrow -Xrf_X - \frac{1}{2} X^2 \sigma^2 f_{XX} - f_t + rf &= 0. \end{aligned} \tag{1.26}$$

□

3. When correctly taking account of the product rule for stochastic variables, it is possible to derive the Black-Scholes equation using any portfolio argument? Yes, this is possible as can be seen from the previous argument, by using the following portfolio

$$(\alpha, Q) = \Big(\frac{B(t)}{f - Xf_X}, -\frac{f_X B(t)}{f - Xf_X}\Big).$$

Finally,

4. How do we create a portfolio to replicate an option/stock? We can't create a self-financing, risk free portfolio that replicates an option/stock since options and stocks are intrinsically risky investments. We may though create a self-financing but risky portfolio to replicate an option. The approach by Øksendal [7] described below demonstrates this approach.

We now take a look at other attempts to derive the Black-Scholes equation and see whether each approach meets the requirements of Proposition 4. For instance, some authors consider a portfolio of a stock and bond to replicate an option, some consider a stock and option to replicate a bond, another considers the idea of a relative portfolio, and another considers a portfolio that by construction replicates a bond or a stock. We would like to be able to see clearly any logical flaws or merits within each and be able to relate all the above approaches.

1.2 Other derivations of the Black-Scholes p.d.e.

1.2.1 Carr's approach

Carr [8] reaches the same conclusions as we have done above. His arguments are very similar to those presented above. He agrees that the traditional portfolio is neither self-financing nor risk free. In the second half of his paper he provides an alternative derivation of the Black-Scholes p.d.e. using the classical portfolio in which $\alpha = -1$ and $Q = f_X$. He claims that we do not need to assume that the portfolio is self-financing and risk free in order to derive the Black-Scholes equation. His approach is not to compute the derivative of the hedged portfolio, but instead to look at the financial 'gain' of the hedged portfolio, which he defines as follows. Consider the following portfolio consisting of one written derivative security $f(t, X)$ and $Q(t, X)$ shares held long, with value process

$$V(t, X) = -f(t, X) + QX.$$

Carr defines the gain as

$$g\Big(V(t, X)\Big) = \int_0^t -df + \int_0^t Q dX$$

Note here again we are ignoring dividend payments for simplicity. Applying Itô's Lemma to compute df and choosing $Q = -\frac{\partial f(t,X)}{\partial X}$ yields

$$g\Big(V(t, X)\Big) = \int_0^t \left[-\frac{\partial f(u, X)}{\partial t} - \frac{1}{2}\sigma^2 X^2 \frac{\partial^2 f(u, X)}{\partial X^2}\right] du.$$

Since this financial gain is deterministic for all time, the absence of arbitrage requires that it be the same as the interest gain for a dynamic position on a

bond B, giving

$$g\Big(B(t)\Big) = \int_0^t r\left(-f(u,X) + X\frac{\partial f(u,X)}{\partial X}\right)du$$

Equating the gains leads to the Black-Scholes p.d.e.. Carr doesn't require that the above portfolio is self-financing, and we know that it is not. The gain function measures only the rise in value of the portfolio due to the change in option value and asset value. By showing that the gains of two hedged, but not necessarily self-financing portfolios are always equal we can derive the Black-Scholes equation. The point to note about this is that the 'gain' function does not incorporate the physical process of hedging and so doesn't tell the whole story.

Carr then asserts that the portfolio given by $\hat{V}(t,X) = \alpha V(t,X)$ is self-financing. We would agree with this statement as in this case we are not holding the number of options constant and the self-financing condition may be satisfied by determining the necessary value of α. Carr does not describe how to determine α. The idea of 'gain' is also used by Davis [9].

1.2.2 Beck's approach

Beck considers a portfolio of stocks and bonds aiming to replicate the value of an option. He requires the portfolio to be self-financing and that its final value must be the same as that of the option. Beck's analysis shows that, in order for a portfolio consisting of stocks and bonds to be self-financing, we require the Black-Scholes equation to hold. The Black-Scholes equation has solution $V = N[d_1]X - e^{-r(T-t)}KN[d_2]$ and so we can think of this as the self-financing requirement. Now, by choice of β, the portfolio value is equivalent (by construction) to the option value. We have $V = f$ and so the value of the

option is given by

$$f = N[d_1]X - e^{-r(T-t)}KN[d_2]. \tag{1.27}$$

This suggests that an option can always be replicated by holding $N[d_1]$ stocks and $KN[d_2]$ bonds. Such a portfolio is automatically self-financing since its form comes from the requirement of self-financing. Based on this result we see that we can create a portfolio consisting of an option and a stock to replicate a bond. By manipulation of the above formula we have

$$e^{-r(T-t)} = B = -\frac{1}{KN[d_2]}f + \frac{N(d_1)}{KN(d_2)}X \tag{1.28}$$

Using the option value formula we compute $N(d_1) = f_X$ and $KN(d_2) = f - f_X X$. Thus, this new portfolio is equivalent to choosing

$$\left(-\frac{1}{f - f_X X}, \frac{f_X}{f - f_X X}, 0 \right)$$

so that this portfolio fits into the regime we derived earlier.

1.2.3 Björk's relative portfolio derivation

The derivation by Björk [5] is a portfolio argument that more naturally incorporates the hedging and self-financing requirements. This is achieved using the idea of a relative portfolio. For example, where Black and Scholes describe the position in the option as the value of the option times the number of options held, Björk describes the position in the option as its value as a fraction of the total value of the portfolio. Let $V(t) = V^h(t)$ be the value process of a portfolio, where $h(t)$ is a vector which denotes the holding in each asset $X(t) = X_0(t), X_1(t), \ldots, X_N(t)$ where $N + 1$ is the number of traded assets in the market. X_0 denotes the risk free bond. Thus a self-financing portfolio is one

in which

$$dV(t) = h(t).dX(t) \tag{1.29}$$

i.e. the portfolio value is given by the dot product of $h(t)$ and $X(t)$

Definition 4. *For a given portfolio h, the corresponding relative portfolio u is given by*

$$u_i(t) = \frac{h_i(t)X_i(t)}{V(t)}, \quad i = 1, \quad \ldots, N. \tag{1.30}$$

Thus

$$\sum_{i=1}^{N} u_i(t) = 1 \tag{1.31}$$

The self-financing condition can now be expressed in terms of the relative portfolio.

Lemma 7. *A portfolio h is self-financing if and only if*

$$dV^h(t) = V^h(t) \sum_{i=1}^{N} u_i(t) \frac{dX_i(t)}{X_i(t)} \tag{1.32}$$

Björk's next result is that a solution of (1.32) can always be realised as the value process of a genuine portfolio.

Lemma 8. *Assume there exists a scalar process Z and a vector process $q = (q_1, \ldots, q_N)$ such that*

$$dZ(t) = Z(t) \sum_{i=1}^{N} q_i(t) \frac{dX_i(t)}{X_i(t)}, \tag{1.33}$$

$$\sum_{i=1}^{N} q_i(t) = 1. \tag{1.34}$$

Now define a portfolio h by

$$h_i(t) = \frac{q_i(t)Z(t)}{X_i(t)}. \tag{1.35}$$

Then the value process V^h is given by $V^h = Z$, h is self-financing, and the corresponding relative portfolio u is given by $u = q$.

Proof. By definition the value process V^h is given by $V^h(t) = h(t)X(t)$, so equations (1.34) and (1.35) give us

$$V^h(t) = \sum_{i=1}^{N} h_i(t)X_i(t) = \sum_{i=1}^{N} q_i(t)Z(t) = Z(t). \tag{1.36}$$

Inserting (1.36) into (1.35) we see that the relative portfolio u corresponding to h is given by $u = q$. Inserting (1.36) and (1.35) into (1.33) we obtain

$$dV^h(t) = \sum_{i=1}^{N} h_i(t)dX_i(t)$$

which shows that h is self-financing. □

We assume as before that the market consists of a stock, an option on that stock, and a bond. The asset price dynamic follows

$$dX(t) = \alpha(t, X(t))X(t)dt + \sigma(t, X(t))X(t)dW_t. \tag{1.37}$$

For simplicity we assume that α and σ are held constant as in the original Black-Scholes model. $W(t)$ is a Wiener process, σ is known as the volatility of $X(t)$ while α is the local mean rate of return. Also available on the market is the risk free asset with price process B. The price process B is the price of a risk free asset if it has the dynamics

$$dB(t) = r(t)B(t)dt$$

where $r(t)$ is the risk-free interest rate. The B-process is therefore given by

$$B(t) = B(0) \exp \int_0^t r(s)ds$$

When $r(t)$ is a deterministic constant one can interpret B as the price of a bond. Now consider the dynamics of the price process of the option (or claim) on the stock $V(t) = f(t, X(t))$ with payoff $F(X(T))$. By Itô's formula we have

$$dV(t) = f_t dt + f_X dX + \frac{1}{2} f_{XX} (dX)^2$$

Substituting for dX from (1.37) we obtain

$$\begin{aligned} dV(t) &= f_t dt + \alpha X f_X dt + \sigma X f_X dW_t + \frac{1}{2}\sigma^2 X^2 f_{XX} dt \\ &= \alpha_V(t) V(t) dt + \sigma_V(t) V(t) dW_t \end{aligned}$$

where the processes $\alpha_V(t)$ and $\sigma_V(t)$ are defined by

$$\alpha_V(t) = \frac{f_t + \alpha X f_X + \frac{1}{2}\sigma^2 X^2 f_{XX}}{f}, \tag{1.38}$$

$$\sigma_V(t) = \frac{\sigma X f_X}{f} \tag{1.39}$$

At this stage we must discuss the idea of arbitrage in greater detail than before.

An arbitrage opportunity on a financial market is a self-financed portfolio h such that

$$V^h(0) = 0$$

$$P(V^h(T) \geq 0) = 1$$

$$P(V^h(T) > 0) > 0$$

An arbitrage possibility allows the investor to make a positive amount of money

out of nothing without taking any risk. This might arise due to mispricing in the market. One of the main assumptions in this derivation is that the market is efficient in the sense that no arbitrage is possible.

Proposition 9. *Suppose that there exists a self-financing portfolio h, such that the value process V^h has the dynamics*

$$dV^h(t) = k(t)V^h(t)dt, \tag{1.40}$$

where k is a continuous adapted process. The no arbitrage assumption implies that $k(t) = r(t)$ for all t.

Proof. Suppose that $r < k$ at some time $t = 0$. Define Δ_t as the time interval between $t = 0$ and the first time at which $r = k$. By continuity of the functions $r(t)$ and $k(t)$ (w.p.1), the interval Δ_t will be well defined and of positive length. During the time interval Δ_t, the interest rate $r(t)$ is strictly less than the growth rate of the value process V^h as denoted above by $k(t)$. Then we can borrow money from the bank at the rate r. This money is immediately invested in the portfolio strategy h where it will grow at the rate k. At the end of this time period sell the portfolio and pay off the bank with this cash. As $k > r$, the cash owed to the bank will be less than the cash value of the portfolio. Thus the net investment will be positive and we have an arbitrage. If on the other hand $r > k$, we short the portfolio h short and invest this money in the bank, and again there is an arbitrage. If we can find a value process of a self-financing portfolio that satisfies the above form, then we know from this proposition that in a no arbitrage market we always will have $k(t) = r(t)$. □

Now form a portfolio based on two assets: the underlying stock and the derivative asset. Denoting the relative portfolio by (u_X, u_V) and using equation

(1.32) we obtain the following dynamics of the value V of the portfolio.

$$dV = Vu_X[\alpha dt + \sigma dW_t] + u_V[\alpha_V dt + \sigma_V dW_t]. \tag{1.41}$$

We now collect the dt and the dW terms to obtain

$$dV = V[u_X\alpha + u_V\alpha_V]dt + V[u_X\sigma + u_V\sigma_V]dW_t \tag{1.42}$$

The only restriction on the relative portfolio is that we must have

$$u_X + u_V = 1,$$

for all t. Let us thus define the relative portfolio by the linear system of equations

$$u_X + u_V = 1, \tag{1.43}$$

$$u_X\sigma + u_V\sigma_V = 0. \tag{1.44}$$

Using this portfolio we see that by its very definition the driving dW term in the V-dynamics of equation (1.42) vanishes completely, leaving us with the equation

$$dV = V[u_X\alpha + u_V\alpha_V]dt. \tag{1.45}$$

We have obtained a riskless portfolio, and because of the requirement that the market is free of arbitrage, we may now use Proposition 9 to deduce that we must have the relation

$$u_X\alpha + u_V\alpha_V = r. \tag{1.46}$$

This is the condition for the absence of arbitrage. It is easily seen that the

system (1.43)-(1.44) has the solution

$$u_X = \frac{\sigma_V}{\sigma_V - \sigma}, \tag{1.47}$$

$$u_V = \frac{-\sigma}{\sigma_V - \sigma} \tag{1.48}$$

which, using (1.39), gives the portfolio more explicitly as

$$u_X = \frac{Xf_X}{Xf_X - f}, \tag{1.49}$$

$$u_V = \frac{-f}{Xf_X - f}. \tag{1.50}$$

Now substitute (1.38), (1.49) and (1.50) into the absence of arbitrage condition (1.46). Then, after some calculations, we obtain the equation

$$f_t + rXf_X + \frac{1}{2}\sigma^2 X^2 f_{XX} - rf = 0. \tag{1.51}$$

Furthermore we must have the relation

$$V(T) = F(X(T)).$$

This portfolio derivation makes an assumption about the dynamics of the underlying portfolio. This assumption appears in equation (1.41) where the dynamics of the relative portfolio are stated to be of this self-financing form. Björk fails to show that this equation holds true.

Approaching the derivation from the point of view of a relative portfolio, it is not necessary to state the actual size of the portfolio. We must remember that the relative portfolio only defines the ratio of the value of each position with regard to the overall value of the portfolio. It is equivalent to not choosing α. Converting between the 'real' position and 'relative' position makes this clear.

Let the real option position in Björk's portfolio be denoted by h_1. Björk gives the (relative) position in the option as

$$\frac{-1}{Xf_X - f} f = \frac{h_1}{V} f.$$

Now consider the position in the option in terms of the framework described earlier in the chapter. The option position is given by α. In Björk's relative portfolio, the position in the option is given by

$$\frac{-1}{Xf_X - f} f = \frac{h_1 f}{V} = \frac{\alpha}{V} f.$$

Now since this is a relative position, we a free to define the overall value of the portfolio as we like while maintaining the ratio of α/V. If we define $\alpha = -\frac{1}{(Xf_X - f)}$ then we have $V = 1$.

The stock position given by Björk (in relative terms) is

$$\frac{f_X}{Xf_X - f} X = \frac{h_2}{V} X.$$

For the same choice of α we have $h_2 = -\alpha f_X$. It is clear that the value of V determines the value of α necessary for self-financing and visa versa. This result is not clear from Björk's derivation and it is not clear for a given value of α (V) how we should choose V (α).

To determine the correct weightings such that a portfolio is guaranteed to be both risk free and self-financing, then the weights on that portfolio must be explicitly of the form described in Proposition 6.

1.2.4 Solving the Black-Scholes equation

So far we have considered derivations of the Black-Scholes equation using simple portfolio arguments. It is instructive to take a more rigorous approach and use results from martingale pricing theory in order to achieve the same results. One very good reference is the book by Øksendal [7]. It is not necessary to present the full derivation here but we will state the results. We will also refer to results about completeness, change of measure and existence of arbitrage as derived in this book. The main result given in Øksendal is that the 'fair' price of the discounted claim is given by the expectation of the future payoff of the claim and can be hedged using the underlying asset and risk free bonds. This relies on the existence of a martingale measure. Such a measure will exist and is unique if the number of sources of randomness is the same as the number of traded underlying assets (completeness), as is the case in this situation.

We know from the arguments in [7] that for a given claim F we can find a hedging portfolio of bonds and the underlying to hedge that portfolio. Using some arbitrage arguments and the uniqueness property of the martingale measure it is possible to conclude that the present value of the claim $f(X,t)$ is given by

$$f(X,t) = E_{\mathbb{Q}}\Big[\xi(t,T)f(X(T))\Big] = \xi(t)E_{\mathbb{Q}}\Big[F(X(T))\Big].$$

where $\mathbb{Q}$ is the equivalent martingale measure and ξ is the numeraire, defined to be $B(T,t)^{-1}$ (where $B(T,T)=1$). We apply the Kolmogorov Backward equation to $f(X(t),T-t) = e^{-r(T-t)}E_{\mathbb{Q}}[F]$ which gives

$$\begin{aligned} -\frac{\partial}{\partial t}e^{-r(T-t)}f &= e^{-r(T-t)}rX\frac{\partial f}{\partial X} + e^{-r(T-t)}\frac{1}{2}\sigma^2X^2\frac{\partial^2 f}{\partial X^2}. \\ \Rightarrow rf - f_t &= rXf_{xx} + \frac{1}{2}\sigma^2X^2f_{xx} \\ \Rightarrow f_t + rXf_{xx} + \frac{1}{2}\sigma^2X^2f_{xx} - rf &= 0 \end{aligned}$$

We also have that $f(F,0) = F[X] = \max(X-K,0)$. We know that in order to hedge this claim we can determine the hedging portfolio $\theta(t)$ given by

$$\xi(t,\omega)(\theta_1(t),\ldots,\theta_n(t))\sigma(t,\omega) = \phi(t,\omega)$$

such that

$$\xi(T)F(\omega) = z + \int_0^T \phi(t,\omega)dW^{\mathbb{Q}}$$

Using a result from Dynkin [10] and Kolmogorov's backward equation the following theorem holds.

Theorem 10. *Let $Y(t)$ be an Itô diffusion in $\mathbb{R}^n$ of the form*

$$dY(t) = b(Y(t))dt + \sigma(Y(t))dW(t), \quad Y(0) = y$$

and $Z(t)$ be the Itô diffusion in $\mathbb{R}^n$ given by

$$dZ(t) = \sigma(Z(t))dW(t), \quad Z(0) = z.$$

Given some restrictions on b and σ as described in [7] and given that we have a complete market we may write

$$h(Y(T)) = E^y_{\mathbb{Q}}\Big[h(Y(T))\Big] + \int_0^T \phi(t,\omega)dW^{\mathbb{Q}},$$

where $\phi = (\phi_1,\ldots,\phi_m)$, with

$$\phi_j(t,\omega) = \sum_{i=1}^{n} \frac{\partial}{\partial y_i}\Big(E^y\Big[h(Z(T-t))\Big]\Big)_{y=Y(t)} \sigma_{ij}(Y(t)), \quad 1 \le j \le m.$$

(Note that in the above notation $E^y[h(Z(t))]$ is the expectation of $h(Z(t))$ subject to $Z(0) = y$).

1.3 Solution using Martingale Pricing Theory

The Black-Scholes market consists of the risk free asset $B(t)$ and one risky asset $X(t)$ whose dynamics are given by

$$\begin{aligned} dX &= \mu X dt + \sigma X dW \\ dB &= rBdt. \end{aligned} \tag{1.52}$$

Since we have the number of traded assets m equals the number of sources of randomness, we have a complete market. The change of measure (where the Girsanov kernel is given by $u(t,\omega)$) exists and is given by the following equation

$$\sigma X(t,\omega)u(t,\omega) = \mu X(t,\omega) - rX(t,\omega)$$

which has the solution

$$u(t,\omega) = \sigma^{-1}(\mu - r).$$

The fair value of a claim with payoff $F(X)$ is given by

$$f(F) = f(X, T-t) = E_{\mathbb{Q}}\Big[\xi(T)F\Big].$$

Under the new measure we have

$$\begin{aligned} dW^{\mathbb{Q}} &= \sigma^{-1}(\mu - r)dt + dW \\ dX &= \mu X dt + \sigma X dW \\ &= \mu X dt + \sigma X(\sigma^{-1}(r-\mu)dt + dW^{\mathbb{Q}}) \\ &= rXdt + \sigma X dW^{\mathbb{Q}} \end{aligned}$$

Now the solution to equation (1.52) is given by

$$X(t) = x_0 \exp\Big(\int_0^t \big(r - \frac{\sigma^2}{2}\big)ds + \int_0^t \sigma dW^{\mathbb{Q}}\Big).$$

At time t, a claim with a maturity date $T \geq t$ is given by

$$\begin{aligned} f(X, T-t) &= \xi(T-t)E_{\mathbb{Q}}\Bigg[f\Big(x_t \exp\Big(\int_t^T \big(r - \frac{\sigma^2}{2}\big)ds + \int_t^T \sigma dW^{\mathbb{Q}}\Big)\Big)\Bigg] \\ &= \xi(T-t)\frac{1}{\sqrt{2V\sigma^2(T-t)}}\int_{\mathbb{R}} f\Big(x_0 \exp\Big[\int_t^T \big(r - \frac{\sigma^2}{2}\big)ds + y\Big]\Big) \\ &\quad \times \exp\Big(-\frac{y^2}{2\sigma^2(T-t)}\Big)dy \end{aligned}$$

since $\int_t^T \sigma dW$ is normally distributed with mean 0 and variance $\sigma^2(T-t)$. In the case of a European Call option we have

$$F(X, T) = \max(X - K, 0)$$

and so the above integral becomes

$$\begin{aligned} f(X, T-t) &= \frac{\xi(T-t)}{\sqrt{2V\sigma^2(T-t)}}\int_{\mathbb{R}} \max\Big(x_0 \exp\Big[\int_t^T \big(r - \frac{\sigma^2}{2}\big)ds + y\Big] - K, 0\Big) \\ &\quad \times \exp\Big(-\frac{y^2}{2\sigma^2(T-t)}\Big)dy \end{aligned}$$

Now since we have constant r and σ we have $\int_t^T \big(r - \frac{\sigma^2}{2}\big)ds = \big(r - \frac{\sigma^2}{2}\big)(T-t)$ and we need only integrate where

$$\begin{aligned} K &\leq x_0 \exp\Big[\big(r - \frac{\sigma^2}{2}\big)(T-t) + y\Big] \\ \Rightarrow \log\Big(\frac{K}{x_0}\Big) - \big(r - \frac{\sigma^2}{2}\big)(T-t) &\leq y. \end{aligned}$$

Therefore we can write our integral as

$$\begin{aligned} f(X,T-t) &= \frac{\xi(T-t)}{\sqrt{2V\sigma^2(T-t)}}\int_{\log\left(\frac{K}{x_0}\right)-\left(r-\frac{\sigma^2}{2}\right)(T-t)}^{\infty}\left(x_0\exp\left[\left(r-\frac{\sigma^2}{2}\right)(T-t)+y\right]-K\right) \\ &\quad\times\exp\left(-\frac{y^2}{2\sigma^2(T-t)}\right)dy \\ &= \frac{\xi(T-t)}{\sqrt{2V\sigma^2(T-t)}} \\ &\quad\times\int_{\log\left(\frac{K}{x_0}\right)-\left(r-\frac{\sigma^2}{2}\right)(T-t)}^{\infty}\left(x_0\exp\left[\left(r-\frac{\sigma^2}{2}\right)(T-t)+y-\frac{y^2}{2\sigma^2(T-t)}\right]\right)dy \\ &\quad-\frac{\xi(T-t)}{\sqrt{2V\sigma^2(T-t)}}\int_{\log\left(\frac{K}{x_0}\right)-\left(r-\frac{\sigma^2}{2}\right)(T-t)}^{\infty}K\exp\left(-\frac{y^2}{2\sigma^2(T-t)}\right)dy \\ &= I_1-I_2 \end{aligned}$$

Let

$$\begin{aligned} y' &= \frac{y}{\sigma\sqrt{T-t}}-\sigma\sqrt{T-t} \\ \Rightarrow dy' &= \frac{1}{\sigma\sqrt{T-t}}dy \end{aligned}$$

and our lower limit becomes

$$\begin{aligned} \frac{\log\left(\frac{K}{x_0}\right)-\left(r-\frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}-\sigma\sqrt{T-t} &= \frac{\log\left(\frac{K}{x_0}\right)-\left(r+\frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} \\ &= -\frac{\log\left(\frac{x_0}{K}\right)+\left(r+\frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}=-d_1 \end{aligned}$$

We rewrite the exponent of the first integral as

$$\left(r-\frac{\sigma^2}{2}\right)(T-t)+y-\frac{y^2}{2\sigma^2(T-t)} = r(T-t)-\frac{y'^2}{2}$$

and so our first integral becomes

$$\begin{aligned} I_1 = x_0 \exp \frac{1}{\sqrt{2V}} \int_{-d_1}^{\infty} \exp\Big[-\frac{y'^2}{2}\Big] dy' &= x_0 \exp \frac{1}{\sqrt{2V}} \int_{-d_1}^{\infty} \exp\Big[-\frac{y'^2}{2}\Big] dy' \\ &= x_0 \exp \frac{1}{\sqrt{2V}} \int_{-\infty}^{d_1} \exp\Big[-\frac{y'^2}{2}\Big] dy' \\ &= x_0 N[d_1] \end{aligned}$$

To evaluate the second integral we instead use the substitution

$$\begin{aligned} y' &= \frac{y}{\sigma\sqrt{T-t}} \\ \Rightarrow dy' &= \frac{1}{\sigma\sqrt{T-t}} dy' \end{aligned}$$

and our lower limit becomes

$$\frac{\log\left(\frac{K}{x_0}\right) - \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}} = -d_2$$

Our second integral becomes

$$\begin{aligned} I_2 &= e^{-r(T-t)} K \frac{1}{\sqrt{2V}} \int_{-d_2}^{\infty} \exp\Big(-\frac{y'^2}{2}\Big) dy' \\ &= e^{-r(T-t)} K \frac{1}{\sqrt{2V}} \int_{-\infty}^{d_2} \exp\Big(-\frac{y'^2}{2}\Big) dy' \\ &= e^{-r(T-t)} K N[d_2] \end{aligned}$$

Combining these two results we have

$$f(X, T-t) = x_0 N[d_1] - e^{-r(T-t)} K N[d_2].$$

1.4 Analysis of the Solution

1.4.1 Implied Volatility

Now, assuming a value for the volatility of the underlying and the interest rate, we can determine the option price for a given strike and time to maturity. Conversely, we should also be able to determine the volatility of an asset if we know the price of the option, strike price, time to maturity and interest rate. This is known as *implied* volatility. It would seem reasonable, given options of different strikes and maturities on the same underlying asset, that the calculated implied volatility would be the same for each (as we are only dealing with one asset). Empirical work shows this not to be the case. Typically the volatility versus strike curve is known as a 'volatility smile', see Figure 1.1. The appearance of the smile may be a result of incorrect assumptions such as the assumption of constant volatility of log-normally distributed returns. To test these assumptions we looked at returns from the Irish Stock Exchange Index, see Figure 1.2 below. We plotted historical volatility and distribution of the returns. There is now widely documented empirical evidence that risky financial asset returns have leptokurtic tails [11]. In the case where the strike price is very high, the call option is deep out of the money and the probability for this option to be exercised is very low. Nevertheless, a leptokurtic right tail will give this option a higher probability of being exercised than a normal distribution would suggest. This higher probability leads to a higher call price and a higher Black-Scholes implied volatility at high strike. Again because of the thicker tail on the left, we expect the probability that an out of the money put option finishes in the money to be higher than that for a normal distribution. Hence the put option price should be greater than that predicted by Black and Scholes. From the Black-Scholes formula we can determine the implied volatility

of a market. If we use Black-Scholes to invert volatility estimates from these option prices, the Black-Scholes implied volatility will be higher than historical volatility. This results in a volatility smile, where implied volatility is much higher at very low and very high strikes. Black and Scholes assume constant volatility in their model. Yet, despite this invalid assumption, the Black-Scholes implied volatility is commonly quoted in the pricing of options.

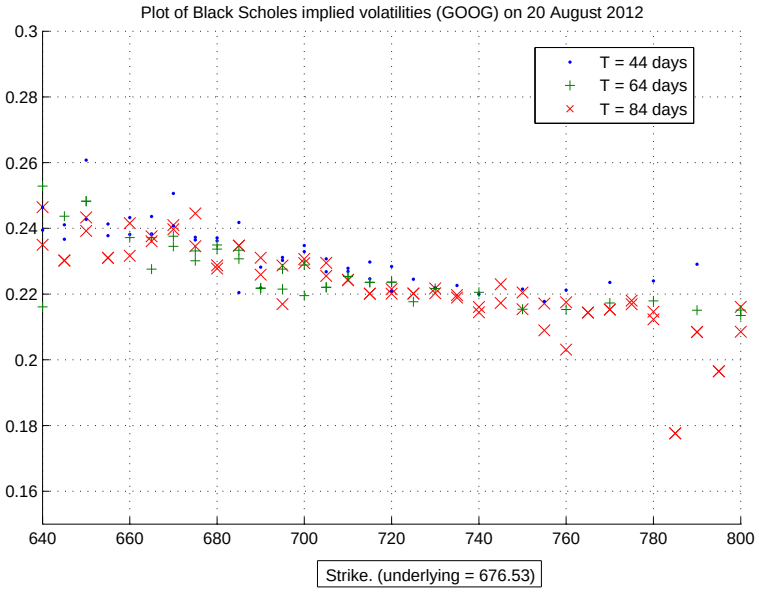

Figure 1.1: Plot of implied volatility for different maturities. We see that the implied volatility depends on maturity and strike price.

1.5 Non-constant volatility models

The 'smile' characteristic of observed implied volatilities was first observed after the stock market crash of October 1987. The effect is evidence of a sudden change in modelling assumptions in which a 'correction' of option prices due to the non-lognormal behaviour of the underlying stock prices was applied. Practitioners had begun to modify the volatility parameter used in the closed-form Black-Scholes equation depending on how close the option was to its at-the-money value and also its time to maturity. In other words, they

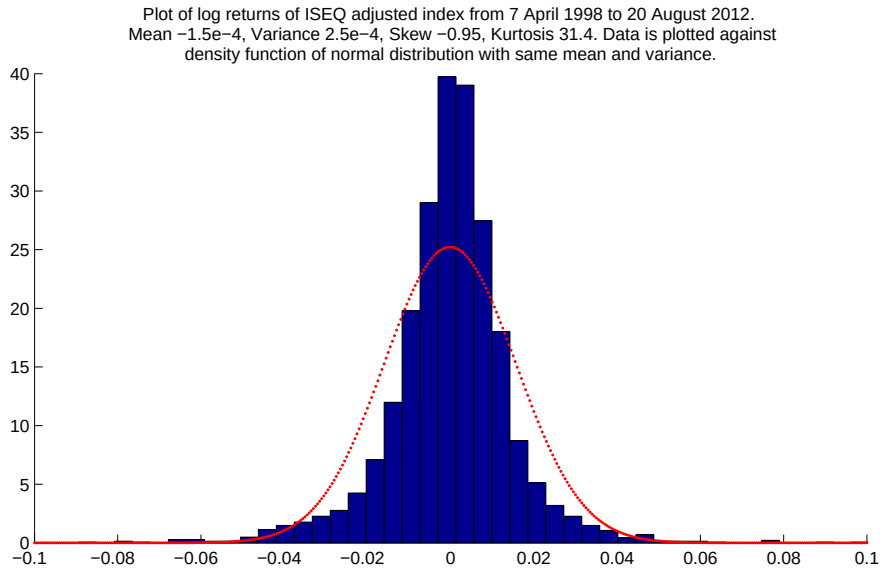

Figure 1.2: Asset returns are assumed to be log-normally distributed. Here we have plotted the log daily returns of the ISEQ index and placed them against a plot of a normal distribution curve of the same mean and sample variance as our data set. We see our marked data has a higher peak and fatter tails.

maintained the use of the Black-Scholes model, but corrected for its constant volatility assumption. It represents the first attempt to relax this assumption. This approach preserves the completeness of the model which, as discussed above, implies a unique fair price for a derivative. It also allows for the construction of hedging strategies.

In another approach, Dupire shows a link between the diffusion process and the implied volatility surface and gives a closed form expression of the surface as a function of market option prices. A difficulty with this approach is that option prices for all strikes and maturities are not always available or reliable. This introduces the need for interpolation and extrapolation techniques which are another source of pricing risk within the model itself.

While the approach of using a volatility surface or smile is theoretically inconsistent, its simplicity, and yet flexibility to match observed prices makes it a very popular approach and is dominant even today. Indeed, more modern pricing models are calibrated such that the volatility implied by those models

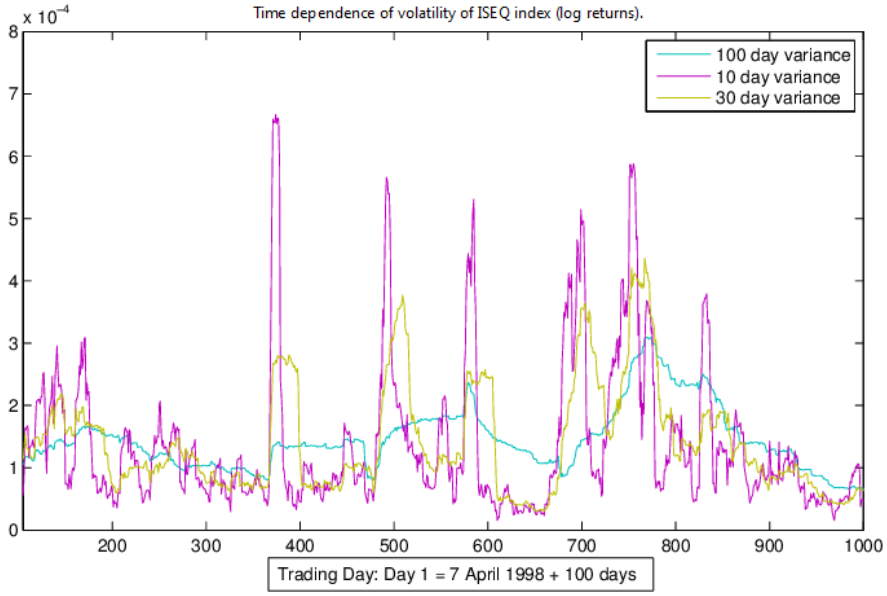

Figure 1.3: Historical Volatility of the ISEQ Index over 1000 days. It is clear that the assumption of constant volatility is incorrect.

matches the implied volatility smile. Such an example is the SABR model [12]. Of course the main use of option pricing models it to price OTC derivatives. The construction of the volatility smile is a method of interpolation, which is then used to price non exchange traded or exotic options. The important considerations in the choice of model is that only the factors which influence the price of the option are taken into account, and that the change in value of the model price reflects changes in those underlying factors. For vanilla European options, the Black-Scholes model meets these requirements. For path dependent and non-vanilla option other models may need to be considered for accurate valuations. For example in path dependant options, or options on a basket of underlying indices, the correlation between volatility and the underlying, or between the set of underlying indices will affect the pricing. In those cases we need to consider a stochastic volatility model. Such models are calibrated to market prices of vanilla options, and the calibrated parameters may be then used in a simulation of the indices if taking a Monte Carlo pricing approach. We will discuss stochastic volatility models in the next section.

Sticking with deterministic volatility models for now, we can consider both discrete time and continuous time examples. We will consider the GARCH and CEV models.

1.5.1 CEV model

This type of model may be described by the s.d.e

$$dX = \mu X dt + \sigma X^{\gamma} dW_t.$$

For $\gamma = 1$ we recover geometric Brownian motion, but for $\gamma > 0$ we see an increased volatility level for higher values of the underlying. We see a similar structure in the the SABR model. A more recent variation of this volatility specification was developed by Rubinstein [13]. Instead of assuming a particular form of the volatility function, Rubinstein's method effectively infers the dependence of volatility on the level of the asset price from traded options at all available strike prices.

1.5.2 GARCH model

A well-known and popular family of models are the ARCH models. ARCH stands for Autoregressive Conditional Heteroscedasticity. It makes use of information on past prices to update the current asset volatility. These models were introduced by Robert Engle (1982) for general time-series modeling. In an ARCH model, the variance that will prevail one step ahead of the current time is a weighted average of past squared asset returns. ARCH can place a greater weight on more recent squared returns than on more distant squared returns. There are many variations on the basic ARCH model. A good survey is given by Poon and Granger [14].

Definition 5. *Let the return of an asset be denoted by r_t. Suppose that the return behaves according to*

$$r_t = \mu + \epsilon_t \tag{1.53}$$

where

$$\epsilon_t = \sqrt{h_t} z_t. \tag{1.54}$$

and z_t are $N[0,1]$ i.i.d random variables. In the ARCH formulation, h_t follows one of the following class of models.

ARCH(q):

$$h_t = \omega + \sum_{k=1}^{q} \alpha_k \epsilon_{t-k}^2 \tag{1.55}$$

where $\alpha_k \geq 0$, $\omega \geq 0$. For finite variance $\sum \alpha_k < 1$.

GARCH(p, q):

$$h_t = \omega + \sum_{k=1}^{q} \alpha_k \epsilon_{t-k}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \tag{1.56}$$

where $\alpha_k \geq 0$, $\beta_k \geq 0$, $\omega \geq 0$. For finite variance $\sum \alpha_j + \sum \beta_k < 1$.

For the results mentioned here on the finite variance conditions see [15] Proposition 3.19. (As is well-known the conditions on α_k, β_k in this definition could be relaxed but we will not go into that here.) Because ARCH models can place greater weight on more recent squared returns than on more distant squared returns, they are able to capture volatility clustering. This refers to the observed tendency of high volatility or low volatility periods to group together. For comparison with stochastic volatility models which will be discussed later, it is important to note that the random source that affects the statistical behavior of returns and volatility through time is the same. As a result, volatility can be estimated directly from the time series of observed returns on an asset. In contrast, the direct estimation of volatility from the returns process is very difficult using stochastic-volatility models. It turns out that there is no easily

computable formula, like the Black-Scholes formula for European option pricing under a GARCH volatility process. We will later return to the GARCH model for a comparison with the Hobson and Rogers model.

1.5.3 Stochastic Volatility Models

A generalisation of constant volatility models such as the Black Scholes model is to assume that the volatility is stochastic, with a noise term which may have some correlation with that of the underlying asset. This is especially important when wanting to price an option with a path dependent payoff. A stochastic volatility model has the ability to reproduce the volatility smile and skew seen in the market. In fact, the GARCH model above is an example of a discrete stochastic volatility model. We will discuss some of the other popular stochastic models in this section. Consider the case of implementing a stochastic volatility model to price an option on a single underlying asset. In the case that the noise term of the volatility differs from that of the underlying asset, the number of random sources is greater than the number of traded assets, and so it is impossible to find a measure under which the discounted underlying asset is a martingale. A risk-free portfolio cannot be created as is done in the Black-Scholes framework. As such no closed form solution is possible. For the purposes of calibration, and subsequent pricing of options, an approximate closed form solution, or a numerical simulation, or both is required. There are approximate closed form solutions to many of the popular stochastic volatility models, and we provide a derivation below of the Heston closed form approximation, taken directly from the book by Gatheral [16].

1.5.4 Heston Model

We take the Heston model as an example of a model with a semi-analytical solution. We give a complete derivation of the semi-analytical solution as the model is very similar to that of the Hobson and Rogers model introduced in Chapter 2. Using the Heston model, a valuation equation, analogous to the Black-Scholes equation, can be derived, the difference being that it takes into account an extra degree of freedom, coming from the dependence on the volatility process. The market is defined by

$$\begin{aligned} dX &= \mu X dt + \sqrt{\nu} X dW_1, \\ d\nu &= -\lambda(\nu - \bar{\nu})dt + \eta\sqrt{\nu} dW_2. \end{aligned}$$

The two increments of Brownian motion, $dW_1$and dW_2 have a correlation of ρ. The value of an option with stochastic volatility is a function of three variables, $f(X, \nu, t)$. We may now attempt to form a hedging portfolio as before in order to replicate a bond. The difference now though is that we have a new source of risk coming from the volatility. To overcome this we need to introduce into our market a second derivative in order to hedge the volatility risk. Denote the value of this derivative by $f_1(X, \nu, t)$. We have that the value process of our hedging portfolio is given by

$$V_t = f - QX - Q_1 f_1,$$

where Q denotes the position in the asset and Q_1 denotes the position an asset which depends on the volatility of the underlying. We have two traded assets and two sources of risk therefore the market is complete and there exists a replicating portfolio for the derivative, which is self-financing. If we assume that we have found the self-financing portfolio which replicates the derivative, then

the change in the value of the portfolio in a time dt is given by

$$\begin{aligned} dV_t &= \left(\frac{\partial f}{\partial t} + \frac{1}{2}\nu X^2 \frac{\partial^2 f}{\partial X^2} + \rho\eta\nu X \frac{\partial^2 f}{\partial X \partial \nu} + \frac{1}{2}\eta^2 \nu \frac{\partial^2 f}{\partial \nu^2}\right) dt \\ &\quad -Q_1 \left(\frac{\partial f_1}{\partial t} + \frac{1}{2}\nu X^2 \frac{\partial^2 f_1}{\partial X^2} + \rho\eta\nu X \frac{\partial^2 f_1}{\partial X \partial \nu} + \frac{1}{2}\eta^2 \nu \frac{\partial^2 f_1}{\partial \nu^2}\right) dt \\ &\quad + \left(\frac{\partial f}{\partial X} - Q_1 \frac{\partial f_1}{\partial X} - Q\right) dX \\ &\quad + \left(\frac{\partial f}{\partial \nu} - Q_1 \frac{\partial f_1}{\partial \nu}\right) d\nu. \end{aligned}$$

To eliminate all randomness from the portfolio we must choose

$$\frac{\partial f}{\partial X} - Q_1 \frac{\partial f_1}{\partial X} - Q = 0$$

to eliminate the dX terms, and

$$\frac{\partial V}{\partial \nu} - Q_1 \frac{\partial V_1}{\partial \nu} = 0$$

to eliminate the $d\nu$ terms. This leaves us with

$$\begin{aligned} dV_t &= \left(\frac{\partial f}{\partial t} + \frac{1}{2}\nu X^2 \frac{\partial^2 f}{\partial X^2} + \rho\eta\nu X \frac{\partial^2 f}{\partial X \partial \nu} + \frac{1}{2}\eta^2 \nu \frac{\partial^2 f}{\partial \nu^2}\right) dt \\ &\qquad -Q_1 \left(\frac{\partial f_1}{\partial t} + \frac{1}{2}\nu X^2 \frac{\partial^2 f_1}{\partial X^2} + \rho\eta\nu X \frac{\partial^2 f_1}{\partial X \partial \nu} + \frac{1}{2}\eta^2 \nu \frac{\partial^2 f_1}{\partial \nu^2}\right) dt \\ &= rV dt \\ &= r(f - QX - Q_1 f_1) dt. \end{aligned} \tag{1.57}$$

Collecting all the f terms on the left hand side, and all the f_1 terms on the right hand side of equation (1.57) we find that

$$\left(\frac{\partial f}{\partial t}+\frac{1}{2}\nu X^2\frac{\partial^2 f}{\partial X^2}+\rho\eta\nu X\frac{\partial^2 f}{\partial X\partial\nu}+\frac{1}{2}\eta^2\nu\frac{\partial^2 f}{\partial\nu^2}-rf+rX\frac{\partial f}{\partial X}\right)\Big/\frac{\partial f}{\partial\nu}$$
$$=\left(\frac{\partial f_1}{\partial t}+\frac{1}{2}\nu X^2\frac{\partial^2 f_1}{\partial X^2}+\rho\eta\nu X\frac{\partial^2 f_1}{\partial X\partial\nu}+\frac{1}{2}\eta^2\nu\frac{\partial^2 f_1}{\partial\nu^2}-rf_1+rX\frac{\partial f_1}{\partial X}\right)\Big/\frac{\partial f_1}{\partial\nu}\quad .$$

Notice that the left hand side of the above equation is a function of f but not f_1 and the right hand side is a function of f_1 but not f_2. Since the two options should have two different payoffs, the only way for this to be possible is for both sides to be independent of the contract type. Both sides can only be functions of the independent variables X, ν and t. Thus we have

$$\frac{\partial f}{\partial t}+\frac{1}{2}\nu X^2\frac{\partial^2 f}{\partial X^2}+\rho\eta\nu X\frac{\partial^2 f}{\partial X\partial\nu}+\frac{1}{2}\eta^2\nu\frac{\partial^2 f}{\partial\nu^2}-rf+rX\frac{\partial f}{\partial X}$$
$$=\quad -\Big(-\lambda(\nu-\bar{\nu})-\phi(X,\nu,t)\Big)\tfrac{\partial f}{\partial\nu}$$

for some function $\phi(X,\nu,t)$ which is known as the market price of risk. Heston makes the assumption that prices process, with the parameters fitted to option prices, generates the risk-neutral measure so the market price of volatility risk ϕ is set to zero.

We now discuss the Heston model in more detail and follow the derivation provided in [16]. We will show a how semi-analytical solution may be found by use of a Fourier Transform. This will be relevant later on when we examine solutions of the Hobson and Rogers model. The valuation equation describing a call option in terms of the Heston model is given by

$$-\frac{\partial f}{\partial\tau}+\frac{1}{2}\nu f_{11}-\frac{1}{2}\nu f_1+\frac{1}{2}\eta^2\nu f_{22}+\rho\eta\nu f_{12}-\lambda(\nu-\bar{\nu})f_2=0 \qquad (1.58)$$

where the subscripts refer to differentiation with respect to x and ν respectively,

$x := \log(Xe^{r\tau}/K), \tau = T - t$. According to Duffie, Pan and Singleton [17] the solution of equation (1.58) has the form

$$K\{e^x P_1(x, \nu, \tau) - P_0(x, \nu, \tau)\},$$

where the first term in the brackets represents the expectation of the final index level given that the option is in the money and the second term represents the probability of exercise. Substituting the proposed solution into equation (1.58) implies that P_0 and P_1 must satisfy the equation

$$\begin{aligned} -\frac{\partial P_j}{\partial \tau} + \frac{1}{2}\nu\frac{\partial^2 P_j}{\partial x^2} - \left(\frac{1}{2} - j\right)\nu\frac{\partial P_j}{\partial x} + \frac{1}{2}\eta^2\nu\frac{\partial^2 P_j}{\partial \nu^2} + \rho\eta\nu\frac{\partial^2 P_j}{\partial x \partial \nu} & \\ +(a - b_j\nu)\frac{\partial P_j}{\partial \nu} &= 0, \end{aligned} \tag{1.59}$$

for $j = 0, 1$, and $\tau > 0$, where

$$a = \lambda\bar{\nu}, b_j = \lambda - j\rho\eta,$$

subject to the terminal condition

$$\begin{aligned} \lim_{\tau \to 0} P_j(x, \nu, \tau) &= \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \\ &=: \theta(x). \end{aligned} \tag{1.60}$$

For any solution P_j of (1.59) and (1.60) define $P_j(x, \nu, 0) = \theta(x)$. We will now solve (1.59) subject to (1.60) using a Fourier transform technique. The Fourier transform of P_j is given by

$$\tilde{P}_j(u, \nu, \tau) = \int_{-\infty}^{\infty} dx e^{-iux} P_j(x, \nu, \tau)$$

Also we have that

$$\tilde{P}_j(u,\nu,0) = \int_{-\infty}^{\infty} dx e^{-iux}\theta(x) = \frac{1}{iu}$$

The inverse transform is given by

$$P_j(x,\nu,\tau) = \int_{-\infty}^{\infty} \frac{du}{2\pi} e^{iux}\tilde{P}_j(u,\nu,\tau) \tag{1.61}$$

Substituting this into equation (1.59) gives

$$-\frac{\partial \tilde{P}_j}{\partial \tau} + \frac{1}{2}u^2\nu\tilde{P}_j - \left(\frac{1}{2} - j\right)u\nu\tilde{P}_j + \frac{1}{2}\eta^2\nu\frac{\partial^2\tilde{P}_j}{\partial\nu^2} + \rho\eta u\nu\frac{\partial\tilde{P}_j}{\partial\nu} + (a - b_j\nu)\frac{\partial\tilde{P}_j}{\partial\nu} = 0 \tag{1.62}$$

Now define

$$\begin{aligned}
\alpha &= -\frac{u^2}{2} - \frac{iu}{2} + iju \\
\beta &= \lambda - \rho\eta j - \rho\eta iu \\
\gamma &= \frac{\eta^2}{2}
\end{aligned}$$

Then equation (1.62) becomes

$$\nu\left\{\alpha\tilde{P}_j - \beta\frac{\partial\tilde{P}_j}{\partial\nu} + \gamma\frac{\partial^2\tilde{P}_j}{\partial\nu^2}\right\} + a\frac{\partial\tilde{P}_j}{\partial\nu} - \frac{\partial\tilde{P}_j}{\partial\tau} = 0 \tag{1.63}$$

Now substitute

$$\begin{aligned}
\tilde{P}_j(u,\nu,\tau) &= \exp\{C_j(u,\tau)\bar{\nu} + D_j(u,\tau)\nu\}\tilde{P}_j(u,\nu,0) \\
&= \frac{1}{iu}\exp\{C_j(u,\tau)\bar{\nu} + D_j(u,\tau)\nu\}
\end{aligned}$$

It follows that

$$\begin{aligned}\frac{\partial \tilde{P}_j}{\partial \tau} &= \left\{\bar{\nu}\frac{\partial C_j}{\partial \tau} + \nu\frac{\partial D_j}{\partial \tau}\right\}\tilde{P}_j \\ \frac{\partial \tilde{P}_j}{\partial \nu} &= D_j\tilde{P}_j \\ \frac{\partial^2 \tilde{P}_j}{\partial \nu^2} &= D_j^2\tilde{P}_j\end{aligned}$$

The equation (1.63) is satisfied if

$$\begin{aligned}\frac{\partial C_j}{\partial \tau} &= \lambda D_j \\ \frac{\partial D_j}{\partial \tau} &= \alpha - \beta D_j + \gamma D_j^2 \\ &= \gamma(D_j - r_+)(D_j - r_-) \qquad (1.64)\end{aligned}$$

where we define

$$r_j^{\pm} = \frac{\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{2\gamma} =: \frac{\beta \pm d}{\eta^2}$$

Note that we have drop the subscript j in the parameters α and β for clarity of notation. Integrating (1.64) with the terminal conditions $C(u,0) = 0$ and $D(u,0) = 0$ gives

$$\begin{aligned}D_j(u,\tau) &= r_j^- \frac{1 - e^{-d\tau}}{1 - ge^{-d\tau}} \\ C_j(u,\tau) &= \lambda\left\{r_j^-\tau - \frac{2}{\eta^2}\log\left(\frac{1 - ge^{-d\tau}}{1 - g}\right)\right\}\end{aligned}$$

where we define

$$g := \frac{r_j^-}{r_j^+}$$

Taking the inverse transform using equation (1.61) and performing the complex integration gives the final form of the probabilities P_j in the form of an integral

of a real valued function.

$$P_j(x,\nu,\tau) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty du\, Re\left\{ \exp\left[\frac{C_j(u,\tau)\bar{\nu} + D_j(u,\tau)\nu + iux}{iu}\right]\right\},$$

where j runs from 0 to 1. This integration may be performed using standard numerical methods.

Chapter 2

The Hobson and Rogers Model

The Black-Scholes model is based on the assumption that the proportional price changes of the asset form a Gaussian process with stationary independent increments. Empirical research has shown that the volatility parameter is not constant [18]. We have seen in Chapter 1 that the constant volatility assumption is inconsistent with the market price of derivatives. Historically, two approaches have been taken to adapt the model. The first is that of 'level dependent volatility' introduced by Cox and Ross (1976) [19]. The basic modelling assumption here is that the volatility is a function of the underlying price of a firm. The second approach is to introduce a second stochastic process for the volatility i.e. 'stochastic volatility'. See, for example, Hull and White [3]. Hobson and Rogers present a model, a special case of which is the level dependent volatility model, but which is also similar to the 'stochastic volatility' approach. This is achieved by making the volatility a function of past returns, as in autoregressive models. It does not require a new source of randomness. Having only one source of randomness is important in option pricing as it allows a unique, preference independent price for the option to be determined. Such models are said to be *complete*. One of the main objectives of the model is to

reproduce the smiles and skews seen in market option prices and this model produces the desired effect. Moreover, since past price movements influence the volatility estimate, if there are large jumps in price, these will be reflected in a large resulting volatility, which mirrors what would be expected to happen in the market.

2.1 Description of Model

The main feature of this model is the specification of volatility in terms of past returns. In this model, the returns, which are used as inputs to the volatility function, are weighted such that more recent returns have a greater influence. An exponential weighting is used as seen in equation (2.6) below. The discounted log-price process is denoted by Z_t, so that

$$Z_t = \log(e^{-rt}P_t),$$

where $P(t)$ is the price process. As in GARCH models, this structure allows for feedback so that shocks in the asset price will be reflected by shocks in volatility. Hobson and Rogers assume that the volatility σ has the form

$$\sigma = \sigma(D_t^{(1)}, ..., D_t^{(n)}).$$

with $D_t^{(m)}$ being given, for $m \in \mathbb{N}$, by

$$D_t^{(m)} = \int_0^\infty \lambda e^{-\lambda u}(Z_t - Z_{t-u})^m du. \tag{2.1}$$

The process $D_t^{(m)}$ is known as the *offset function* of order m. The parameter λ determines the rate at which past data is discounted. Also we see that the

difference

$$\begin{aligned} Z_t - Z_{t-u} &= \log\left(\tfrac{e^{-rt}P_t}{e^{-r(t-u)}P_{t-u}}\right) \\ &= \log\left(e^{-ru}\tfrac{P_t}{P_{t-u}}\right) \end{aligned}$$

can interpreted as the return on the asset between some previous time and today. The model then assumes that

$$dZ_t = \mu(D_t^{(1)}, ..., D_t^{(n)})dt + \sigma(D_t^{(1)}, ..., D_t^{(n)})dW_t^{\mathbb{P}} \tag{2.2}$$

where $W_t^{\mathbb{P}}$ is some $\mathbb{P}$-Wiener process. We see in [1] that in the case $n = 1$ we have

$$dD_t = dZ_t - \lambda D_t dt, \tag{2.3}$$

where $D_t = D_t^{(1)}$, giving

$$dD_t = (\mu(D_t) - \lambda D_t)dt + \sigma(D_t)dW_t^{\mathbb{P}}. \tag{2.4}$$

Choosing $n = 1$ means that the volatility is a function of the first offset alone. The motivation behind this choice is due to the simplicity it brings to the calculation, but also that if we take $\left(D_t^{(1)}\right)^2 = \left(Z_t - \int_0^\infty \lambda e^{-\lambda u} Z_{t-u} du\right)^2$ we see that this acts like a variance. We can think of the term on the right as a weighted exponential mean, so that the offset function of order 1 gives us the difference between today's asset price, Z_t, and the long run mean asset price. We aim now to follow the arguments in [1] and derive a partial differential equation for the price of a European Call option using a martingale approach. It is assumed that the option price f is, as usual, a function of the underlying price process and time and also a function of the first offset function. This assumption will be discussed later on. The partial differential equation will be derived by looking at the Itô operator applied to $f(P_t, D_t, t)$. By using a change

of measure we can use our assumption of no arbitrage to write down the p.d.e. from this expression. Using (2.3) we first have

$$\log(e^{-rt}P_t) \equiv Z_t = Z_0 + (D_t - D_0) + \lambda \int_0^t D_u \, du.$$

We now consider a change of measure which will show that the price process $e^{-rt}P_t$ is in fact a martingale relative to this measure. We will use the Girsanov Theorem.

Theorem 11. *Let $W_t^{\mathbb{P}}$, $t > 0$, be a d-dimensional standard $\mathbb{P}$-Wiener process on $(\Sigma, \mathcal{F}, \mathbb{P})$ and let ϕ be any d-dimensional adapted column vector process. Choose a fixed T and define the process L on $[0, T]$ by*

$$dL_t = \varphi_t L_t dW_t^{\mathbb{P}}$$
$$L_0 = 1,$$

i.e.

$$L_t = \exp\left[\int_0^t \varphi_s dW_s^{\mathbb{P}} - \frac{1}{2}\int_0^t \|\varphi_s\|^2 ds\right]$$

Assume that

$$E^{\mathbb{P}}[L_T] = 1,$$

and define the new probability measure $\mathbb{Q}$ on $\mathcal{F}_T$ by

$$L_T = \frac{d\mathbb{Q}}{d\mathbb{P}}, \quad \text{on } \mathcal{F}_T.$$

Then

$$dW_t^{\mathbb{P}} = \phi_t dt + dW_t^{\mathbb{Q}},$$

where $W^{\mathbb{Q}}$ is a $\mathbb{Q}$-Wiener process.

We now let $\varphi_t = -\frac{1}{2}\sigma(D_t) - \mu(D_t)/\sigma(D_t)$ and consider the process

$W_t^{\mathbb{Q}} \equiv W_t^{\mathbb{P}} - \int_0^t \varphi_t du$. Define a new measure $\mathbb{Q}$ on $\mathcal{F}_T$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\Big[\int_0^t \varphi_t dW_u^{\mathbb{P}} - \frac{1}{2}\int_0^t \|\varphi_t\|^2 du\Big]$$

We can now rewrite (2.3) as

$$\begin{aligned} dD_t &= (\mu(D_t) - \lambda D_t)dt + \sigma(D_t)(\varphi(S_t) + dW_t^{\mathbb{Q}}) \\ &= (\mu(D_t) - \lambda D_t)dt + \sigma(D_t)\left(-\frac{1}{2}\sigma(D_t) - \mu(D_t)/\sigma(D_t))dt + dW_t^{\mathbb{Q}}\right) \\ &= -\left(\frac{1}{2}\sigma(D_t)^2 + \lambda D_t\right)dt + \sigma(D_t)dW_t^{\mathbb{Q}}. \end{aligned}$$

where $W^{\mathbb{Q}}$ is a $\mathbb{Q}$-Brownian motion. We note that

$$\begin{aligned} d(e^{-rt}P_t) \equiv d(e^{Z_t}) &= e^{Z_t}dZ_t + \frac{1}{2}e^{Z_t}\sigma(D_t)^2(dZ_t)^2 \\ &= e^{Z_t}\left(-\frac{\sigma(D_t)^2}{2}dt + \sigma(D_t)dW^{\mathbb{Q}} + \frac{\sigma(D_t)^2}{2}dt\right) \end{aligned}$$

where we have written dZ in terms of our new measure. We see now that the dt terms cancel and thus

$$d(e^{-rt}P_t) = P_t e^{-rt}\sigma dW^{\mathbb{Q}}.$$

The absence of the drift terms shows us that $e^{-rt}P_t$ is a martingale under $\mathbb{Q}$ so that $\mathbb{Q}$ is an equivalent martingale measure. The equivalence property is discussed in Appendix A of [1]. We may also write the above in terms of dP_t since

$$d(e^{-rt}P_t) = e^{-rt}(-rP_t dt + dP_t)$$

and so

$$dP_t = rP_t dt + \sigma(D_t)P_t dW^{\mathbb{Q}}.$$

Proposition 12. *If we assume the process $Z_t = \log(e^{-rt}P_t)$ follows the*

following process,

$$dZ_t = \mu(D_t^{(1)}) + \sigma(D_t^{(1)})dW^{\mathbb{P}}$$

where $\mathbb{P}$ is some $\mathbb{P}$-Brownian motion, then the price process P_t and the offset process $D_t^{(1)}$ obey the following stochastic differential equations under the risk neutral measure $\mathbb{Q}$, namely

$$\begin{aligned} dP_t &= rP_t dt + \sigma(D_t)P_t dW^{\mathbb{Q}}, \\ dD_t &= -\left(\frac{1}{2}\sigma(D_t)^2 + \lambda D_t\right)dt + \sigma(D_t)dW^{\mathbb{Q}}, \end{aligned}$$

where $W^{\mathbb{Q}}$ is a $\mathbb{Q}$-Brownian motion.

Next we introduce the theory of martingale pricing. We consider the primary market to consist of the bank account and the underlying price process. The task is to determine a `reasonable` price process $f(P_t, D_t, T-t)$ and we assume that the market is arbitrage free. The choice of $T-t$ as the time variable comes from [1]. Also, our data consists of option prices with varying maturities. The third parameter of the function f is the time to maturity of the option. The function does not describe how the option price changes with time, it describes how the option prices changes as a function of maturity dates. This may seem a simple point to make but it is one that is easily overlooked and may otherwise be a source of confusion later on. The derivative should be priced in a way that is consistent with the prices of the underlying assets. More precisely, the extended market (the market with the derivative price process included) should also be free of arbitrage possibilities. This requirement is equivalent, under the first fundamental theorem of finance [20], that all price processes in the market are martingales under some martingale measure. In the above calculations we have found a measure under which the price process is a martingale. It is shown in [5] that this measure is unique and so by the

definition of a martingale measure we obtain

$$\frac{f(P_t, D_t, T-t)}{B_t} = E^{\mathbb{Q}}\left[\frac{f(P_T, D_T, 0)}{B_T}\middle|\mathcal{F}_t\right] = E^{\mathbb{Q}}\left[\frac{\Phi(P_T)}{B_T}\middle|\mathcal{F}_t\right]$$

where B_t is the bank account and $\Phi(X) = \max\{X - K, 0\}$ where K is the strike price of the European option. We thus have the following result.

Theorem 13. *(Risk Neutral Valuation Formula) Assuming the existence of a short rate, the arbitrage free price process for the claim f is given by*

$$f(P_t, D_t, T-t) = B_t E^{\mathbb{Q}}\left[\frac{\Phi(P_T)}{B_T}\middle|\mathcal{F}_t\right]$$

where $\mathbb{Q}$ is the unique martingale measure.

We also have

$$B_t = B_0 \exp\left[\int_0^t r(s)ds\right]$$

and so

$$\begin{aligned} f(P_t, D_t, T-t) &= E^{\mathbb{Q}}\left[e^{-\int_t^T r(s)ds} \cdot \Phi(P_T)\middle|\mathcal{F}_t\right] \\ &= e^{-r(T-t)} E^{\mathbb{Q}}\left[\Phi(P_T)\middle|\mathcal{F}_t\right] \end{aligned}$$

where in the last line we assumed that $r(t)$ is a constant function. By the Feynman-Kac formula [21] (Karatzas and Shreve 1988) $f(P_t, D_t, T-t)$ satisfies the following partial differential equation

$$rP_t f_P - rf - \lambda D_t f_D - f_t + \frac{\sigma(D_t)^2}{2}\left(-f_D + P_t^2 f_{PP} + f_{DD} + 2P_t f_{PD}\right) = 0$$

Choosing $T - t$ as the time variable in f resulting in a negative sign in front of the time derivative. This is all consistent with [1]. We can, of course, see how the above equation simplifies to the Black-Scholes equation if we remove the

dependence on D_t and reverse the sign on the time derivative.

Proposition 14. *Under the Hobson and Rogers model, the price of a European call option with strike price K and maturity T, with an underlying risk free rate r obeys the following partial differential equation*

$$rP_t f_P - rf - \lambda D_t f_D - f_t + \frac{\sigma(D_t)^2}{2}\Big(- f_D + P_t^2 f_{PP} + f_{DD} + 2P_t f_{PD}\Big) = 0 \quad (2.5)$$

with boundary conditions

$$f(P_T, D_T, 0) = \max(P_T - K, 0)$$

where P_t and D_t are the price process and first offset process respectively.

A transformed version of equation (2.5) has been solved numerically by Foschi and Pascucci [22]. This transformation is outlined in the next section. Hobson and Rogers show that taking $\sigma(D_t) = \eta\sqrt{1 + \epsilon(D_t)^2}$ as the specification of the volatility function is enough to reproduce the smiles in the Black-Scholes implied volatility. So far we have only considered the case $n = 1$, in which

$$D_t^{(1)} = D_t = \int_0^\infty \lambda e^{-\lambda u}(Z_t - Z_{t-u})\, du \quad (2.6)$$

Thinking of this integral as a weighted sum of 'returns' we see that there is the possibility that negative returns will be cancelled by positive ones. This represents a loss of potentially important information. Choosing $m = 2$ removes this possibility. We now derive the partial differential equation where we assume our option $f = f(P_t, D_t^{(2)}, T - t)$ i.e. is a function of the second offset only. The general formula for $dD_t^{(m)}$ is derived by Hobson and Rogers. We have

$$dD_t^{(m)} = mD_t^{(m-1)}dZ_t + \frac{m(m-1)}{2}D_t^{(m-2)}d\langle Z\rangle_t - \lambda D_t^{(m)}dt$$

and so for $m = 2$ we have

$$dD_t^{(2)} = 2D_t^{(1)} dZ_t + d\langle Z \rangle_t - \lambda D_t^{(2)} dt.$$

Now we substitute for dZ_t and use the change of measure as before to find

$$dD_t^{(2)} = \left(\sigma^2(1 - D_t^{(1)}) - \lambda D_t^{(2)}\right) dt + 2\sigma D_t^{(1)} dW^{\mathbb{Q}}$$

where $\sigma = \sigma(D_t^{(2)})$.

Proposition 15. *The Feynman-Kac formula for $f = f(P_t, D_t^{(2)}, T - t)$ now yields*

$$\begin{aligned} \sigma^2 \left((1 - D_t^{(1)}) f_{D^{(2)}} + (D_t^{(1)})^2 f_{D^{(2)}D^{(2)}} + \frac{1}{2} P_t^2 f_{PP} + 2D_t^{(1)} f_{D^{(2)}P}\right) & \\ - f_t + r P_t f_P - \lambda D_t^{(2)} f_{D^{(2)}} - rf &= 0 \end{aligned}$$

Proof. We have as usual that the fair price of a claim F is given by $f(F, T - t) = e^{-r(T-t)} E^{\mathbb{Q}}[F]$ and so by the Kolmogorov Backward equation we can write

$$-\frac{\partial f}{\partial t} = \mu(X)\frac{\partial f}{\partial X} + \mu(D)\frac{\partial f}{\partial D} + \frac{1}{2}\sigma(X)^2 \frac{\partial^2 f}{\partial X^2} + \frac{1}{2}\sigma(D)^2 \frac{\partial^2 f}{\partial D^2} + \sigma(X)\sigma(D)\frac{\partial^2 f}{\partial X \partial D}$$

as required, where $\mu(X)$, $\mu(D)$, $\sigma(X)$ and $\sigma(D)$ are defined by

$$\begin{aligned} dX &= \mu(X)dt + \sigma(X)dW^{\mathbb{Q}} \\ &= rXdt + \sigma X dW^{\mathbb{Q}} \end{aligned}$$

and

$$\begin{aligned} dD = dD^{(2)} &= \mu(D)dt + \sigma(D)dW^{\mathbb{Q}} && (2.7) \\ &= \left(\sigma^2(1 - D_t^{(1)}) - \lambda D_t^{(2)}\right) dt + 2\sigma D_t^{(1)} dW^{\mathbb{Q}}. \quad \square && (2.8) \end{aligned}$$

2.2 Transformation of the Hobson and Rogers p.d.e.

Proposition 16. *Using a transformation of variables given by*

$$x = \log(\frac{P}{K}) - r(T-t), \tag{2.9}$$

$$y = \log(\frac{P}{K}) - r(T-t) - D, \quad \textit{and} \tag{2.10}$$

$$\tau = -\lambda(T-t), \tag{2.11}$$

then

$$f(P, D, T-t) \to Ke^{r(T-t)}V(x, y, \tau), \tag{2.12}$$

the partial differential equation describing the price of a European call option as given in Proposition 14 may be written as

$$\frac{\sigma(D_t)^2}{2\lambda}(V_{xx} - V_x) + (x-y)V_y - V_t = 0, \tag{2.13}$$

with the boundary condition

$$V(x_T, y_T, 0) = (e^{x_T} - 1)^+. \tag{2.14}$$

Proof. The Hobson and Rogers p.d.e. for the case of the first offset function is given in Proposition 14 by

$$rP_t f_P - rf - \lambda D_t f_D - f_t + \frac{\sigma(D_t)^2}{2}\left(-f_D + P_t^2 f_{PP} + f_{DD} + 2P_t f_{DP}\right) = 0, \tag{2.15}$$

from equation (2.5). Using the stated transformation of variables we may

rewrite the above partial derivatives as

$$\begin{aligned}
f_P &= e^{r(T-t)}K\left(\frac{V_x}{P}+\frac{V_y}{P}\right),\\
f_{PP} &= e^{r(T-t)}K\left(\frac{1}{P^2}(V_{xx}+V_{xy}-V_x)+\frac{1}{P^2}(V_{yy}+V_{yx}-V_y)\right),\\
f_{DP} &= e^{r(T-t)}K\left(-\frac{1}{P}V_{xy}-\frac{1}{P}V_{yy}\right),\\
f_D &= -e^{r(T-t)}KV_y,\\
f_{DD} &= e^{r(T-t)}KV_{yy},\\
f_t &= e^{r(T-t)}K\left(-rV+rV_x+rV_y+\lambda V_t\right).
\end{aligned}$$

We substitute these partial derivatives into equation (2.15), and using the identity $x-y\equiv D$ we obtain

$$\begin{aligned}
&re^{-rt}(V_x+V_y)-re^{-rt}V+\lambda(x-y)e^{-rt}V_y+re^{-rt}V+e^{-rt}(-rV_x-rV_y-\lambda V_t)\\
&+\frac{e^{-rt}\sigma(D_t)^2}{2}\left(V_y-V_x+V_{xx}+V_{xy}+V_{yx}+V_{yy}-V_y+V_{yy}-2V_{xy}-2V_{yy}\right)=0,
\end{aligned}\tag{2.16}$$

which simplifies to

$$\frac{\sigma(x-y)^2}{2\lambda}(V_{xx}-V_x)+(x-y)V_y-V_t=0.\tag{2.17}$$

Now for the boundary conditions. Since $f(P_T,D_T,0)=\max\{P_T-K,0\}$, we have $f(P_T,D_T,0)=KV(\log(\frac{P}{K}),\log(\frac{P}{K})-D,0)$. Therefore $(P_T-K)^+$ becomes $(Ke^x-K)^+$ and

$$KV(\log(\frac{P}{K}),\log(\frac{P}{K})-D,0)=K(e^x-1)^+,$$

which leads to the final condition

$$V(x_T,y_T,0)=(e^{x_T}-1)^+.$$ □

2.3 Relationship with GARCH models

We would like to make a direct comparison between the Hobson and Rogers model and the standard autoregressive models such as ARCH and GARCH. The idea is that if we can fit a GARCH model to a given data set and we know approximately the relationship between GARCH and the Hobson and Rogers model, we can then approximately calibrate the Hobson and Rogers model from the calibrated GARCH model. In fact Hobson and Rogers refer directly to the ARCH family of models in motivating the form of the offset function. The results of this calibration should give us some idea of what to expect in more accurate calibration procedures. We aim to write the volatility specification of each model in terms of variance. Discretising the Hobson and Rogers model leads to

$$dZ_t \cong Z_{t+1} - Z_t = \mu(D_t^{(1)}, ..., D_t^{(n)})dt + \sigma_t(D_t^{(1)}, ..., D_t^{(n)})z_t.$$

where the z_t are $N[0, \sqrt{dt}]$, i.i.d. random variables. Note that

$$Z_{t+1} - Z_t = \log\left(\frac{P_{t+1}}{P_t}\right) \simeq r_t + \mathcal{O}(r_t^2),$$

where

$$r_t = \frac{P_{t+1} - P_t}{P_t}. \tag{2.18}$$

For the moment we take t to be measured in days, and σ_t to be the daily volatility. This conforms to the notation used in autoregressive models. We will return to the issue of time scaling in grater detail in a later section. For simplicity, we assume in both models that the drift μ is zero. Thus

$$r_t = \sigma_t(D_t^{(1)}, ..., D_t^{(n)})z_t$$

The general form of the autoregressive models with zero drift is given by

$$r_t = \sqrt{h_t} z_t = \epsilon_t \tag{2.19}$$

and

$$h_t = \omega + \sum_{k=1}^{q} \alpha_k \epsilon_{t-k}^2 + \sum_{j=1}^{p} \beta_j h_{t-j}$$

where h_t is the conditional volatility and z_t is the Brownian motion increment. For more details see Definition 5, Chapter 1. We can clearly see that the σ term in the Hobson and Rogers model 'looks like' the $\sqrt{h_t}$ term of the GARCH model. Firstly, we find the maximum likelihood parameters (ω, α, β) of the GARCH$(1,1)$ model. We may then obtain an estimate of the parameters of the Hobson and Rogers model by matching the expected values of the GARCH and Hobson and Rogers volatility processes. We can expand the σ^2 term of the Hobson and Rogers model and the h_t term of the GARCH(1,1) model in terms of ϵ_{t-i} for all i and take the expectation of each process.

First dealing with GARCH we must substitute our definition of h_t into h_{t-j} leaving us with a new ϵ term and another h_{t-j} term. With each substitution we move further back in time and may go as far back as we wish. Thus,

$$\begin{aligned}
h_t &= \omega + \alpha\epsilon_{t-1}^2 + \beta h_{t-1} \\
&= \omega + \alpha\epsilon_{t-1}^2 + \beta(\omega + \alpha\epsilon_{t-2}^2 + \beta h_{t-2}) \\
&= \omega + \alpha\epsilon_{t-1}^2 + \beta\Big(\omega + \alpha\epsilon_{t-2}^2 + \beta(\omega + \alpha\epsilon_{t-3}^2 + \beta h_{t-3})\Big) \\
&\vdots \\
&= \omega(1 + \sum_{i=1}^{\infty} \beta^i) + \alpha \sum_{i=1}^{\infty} \beta^{i-1} \epsilon_{t-i}^2.
\end{aligned}$$

Taking the expected value, and noting that $E[\epsilon_t^2] = h_t$ we have

$$E[h_t] = \omega(1 + \sum_{i=1}^{\infty} \beta^i) + \alpha \sum_{i=1}^{\infty} \beta^{i-1} h_{t-1} \tag{2.20}$$

Now we have expressed the volatility in terms of each h_t which will be used later on.

The volatility specification of the Hobson and Rogers model i.e. σ_{HR}, is given by

$$\sigma_{HR}^2 = a^2 + b^2 D_t^2, \tag{2.21}$$

where we substitute our definition of the offset function D_t from equation (2.6). We must now discretise the offset function which will allow us to make the term by term comparison with GARCH. We apply the GARCH model to daily closing prices as described later. Before we discretise the offset function we consider the time scaling of the parameters. In the case that we choose to use annual, daily or inter day data, we should scale λ appropriately.

Proposition 17. *A time-scaled offset function, with time scaled by a factor k, is given by*

$$\hat{D_r}^{(m)} = \int_0^{\infty} \hat{\lambda} e^{-\hat{\lambda} v} (\hat{Z}_r - \hat{Z}_{r-v})^m dv \tag{2.22}$$

where $\hat{\lambda} = k\lambda$, $\hat{Z}_t = Z_{kt}$ and $v(u) = ku$.

Proof. Recall that

$$D_t^{(m)} = \int_0^{\infty} \lambda e^{-\lambda u} (Z_t - Z_{t-u})^m du, \tag{2.23}$$

with the units of u being years. If we let $v(u) = u/k$, where k is the scaling constant, then

$$k dv(u) = du$$

and we can rewrite (2.23) in terms of our new variable as

$$\begin{aligned} D_t^{(m)} &= \int_0^\infty \lambda e^{-\lambda v k}(Z_t - Z_{t-(vk)})^m k v'(u) du \\ &= \int_0^\infty \hat{\lambda} e^{-\hat{\lambda} v}(Z_t - Z_{t-(kv)})^m dv \end{aligned} \tag{2.24}$$

where $\hat{\lambda} = k\lambda$. Now, for simplicity of notation we define $\hat{Z}_t = Z_{kt}$ and $\hat{D}_t^{(m)} = D_{kt}^{(m)}$. Then (2.24) becomes

$$\hat{D}_t^{(m)} = \int_0^\infty \hat{\lambda} e^{-\hat{\lambda} v}(\hat{Z}_t - \hat{Z}_{t-v})^m dv \qquad \square \tag{2.25}$$

We wish to work in units of days so we will choose $k = 1/252$. Discretising our new offset function we have

$$D_{t+1} = \sum_{i=1}^{\infty} \hat{\lambda} e^{-\hat{\lambda} \sum_{j=t-i}^{t} \Delta_j}(Z_t - Z_{t-i})\Delta_i,$$

where Δ_i is the time difference between successive quotes. Since we are working only with end-of-day data, we have $\Delta_j = 1$ (day) for all j. The above equation becomes

$$D_{t+1} = \sum_{i=1}^{\infty} \hat{\lambda} e^{-\hat{\lambda} i}(Z_t - Z_{t-i}).$$

Here we are using the first offset function. Given the value of the underlying up to time t we can now calculate the offset at time $t+1$. Notice that $Z_t - Z_{t-i}$ can be written as

$$\begin{aligned} Z_t - Z_{t-i} = \log\left(\frac{P_t}{P_{t-i}}\right) &= \log\left(\frac{P_t}{P_{t-1}}\right) + \log\left(\frac{P_{t-1}}{P_{t-2}}\right) + \ldots + \log\left(\frac{P_{t-(i-1)}}{P_{t-i}}\right) \\ &\simeq r_{t-1} + r_{t-2} + \ldots + r_{t-i} \\ &= \sum_{k=1}^{i} \epsilon_{t-k}. \end{aligned}$$

Substituting the above equation into our discretised offset function and the

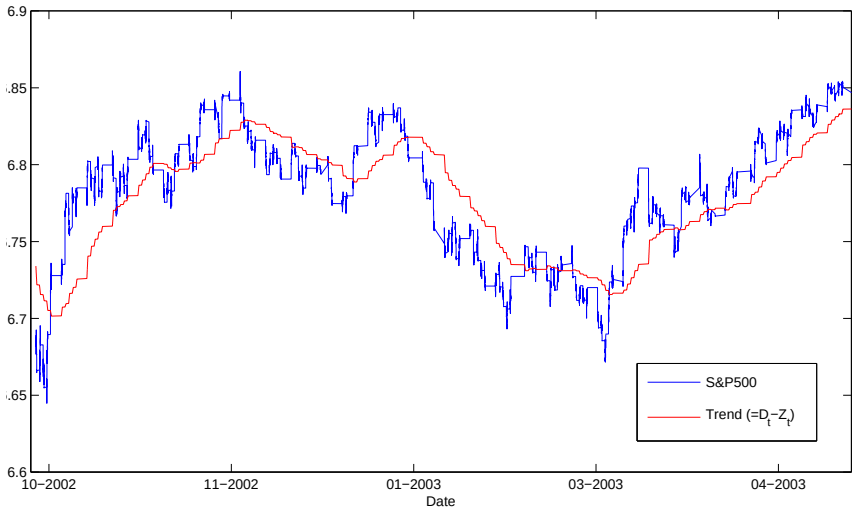

Figure 2.1: Plot of log asset price vs. $\sum_{i=1}^{20} \hat{\lambda} e^{-\hat{\lambda}} Z_{t-i}$

offset function into σ_{HR} as given by equation (2.21) we have

$$\sigma^2_{t+1,HR} = a^2 + b^2 \left(\sum_{i=1}^{\infty} \hat{\lambda} e^{-\hat{\lambda}} \sum_{k=1}^{i} \epsilon_{t-k} \right)^2 . \tag{2.26}$$

We may simplify this expression by interchanging the order of the double summation and gathering together the coefficients for each ϵ term. We write the double summation as

$$\begin{aligned}
\sum_{i=1}^{\infty} \hat{\lambda} e^{-\hat{\lambda} i} \sum_{k=1}^{i} \epsilon_{t-k} &= \hat{\lambda} \Big\{ \epsilon_{t-1}(e^{-\hat{\lambda}} + e^{-2\hat{\lambda}} + \ldots) \\
&\quad + \epsilon_{t-2}(e^{-2\hat{\lambda}} + e^{-3\hat{\lambda}} + \ldots) \\
&\quad \vdots \\
&\quad + \epsilon_{t-i}(e^{-i\hat{\lambda}} + e^{-(i+1)\hat{\lambda}} + \ldots) + \ldots \Big\} \\
&= \hat{\lambda} \Big\{ \epsilon_{t-1} \frac{e^{-\hat{\lambda}}}{1 - e^{-\hat{\lambda}}} + \epsilon_{t-2} \frac{e^{-2\hat{\lambda}}}{1 - e^{-\hat{\lambda}}} + \ldots + \epsilon_{t-i} \frac{e^{-i\hat{\lambda}}}{1 - e^{-\hat{\lambda}}} + \ldots \Big\} \\
&= \frac{\hat{\lambda}}{1 - e^{-\hat{\lambda}}} \sum_{j=1}^{\infty} \epsilon_{t-j} e^{-j\hat{\lambda}}
\end{aligned}$$

The double summation has now become

$$\frac{\hat{\lambda}}{1-e^{-\hat{\lambda}}}\sum_{j=1}^{\infty}\epsilon_{t-j}e^{-j\hat{\lambda}}.$$

We substitute this back into our expression (2.26) for $\sigma^2_{t+1,HR}$ to obtain

$$\sigma^2_{t+1,HR} = a^2 + b^2\left(\frac{\hat{\lambda}}{1-e^{-\hat{\lambda}}}\sum_{j=1}^{\infty}\epsilon_{t-j}e^{-j\hat{\lambda}}\right)^2.$$

Then

$$\begin{aligned}
\mathbb{E}[\sigma^2_{t+1,HR}] &= a^2 + \mathbb{E}\left[\left(c\sum_{j=1}^{\infty}\epsilon_{t-j}e^{-j\hat{\lambda}}\right)^2\right] \\
&= a^2 + \mathbb{E}\Big[c\Big(\epsilon^2_{t-1}e^{-2(1)\hat{\lambda}} \\
&\qquad +2\epsilon_{t-1}\epsilon_{t-2}e^{-3\hat{\lambda}} + \epsilon^2_{t-2}e^{-2(2)\hat{\lambda}} \\
&\qquad +2\epsilon_{t-1}\epsilon_{t-3}e^{-3\hat{\lambda}} + 2\epsilon_{t-2}\epsilon_{t-3}e^{-5\hat{\lambda}} + \epsilon^2_{t-3}e^{-2(3)\hat{\lambda}} \\
&\qquad +\dots\Big)\Big]
\end{aligned}$$

where

$$c = \frac{b^2\hat{\lambda}^2}{(1-e^{-\hat{\lambda}})^2}.$$

Taking the expectation of each of the above terms we have

$$\mathbb{E}[\sigma^2_{t+1,HR}] = a^2 + c\Big(h_{t-1}e^{-2(1)\hat{\lambda}} + h_{t-2}e^{-2(2)\hat{\lambda}} + h_{t-3}e^{-2(3)\hat{\lambda}} + \dots\Big). \qquad (2.27)$$

Here we have made use of $\mathbb{E}[\epsilon_t^2] = h_t$ and

$$\begin{aligned}
\mathbb{E}[\epsilon_t\epsilon_{t-1}] &= \mathbb{E}\Big[\mathbb{E}[\epsilon_t\epsilon_{t-1}|\mathcal{F}_t]\Big] \\
&= \mathbb{E}\Big[\epsilon_{t-1}\mathbb{E}[\epsilon_t|\mathcal{F}_t]\Big] \\
&= \mathbb{E}\Big[\sqrt{h_{t-1}}z_{t-1}\mathbb{E}[\epsilon_t|\mathcal{F}_t]\Big] \\
&= \mathbb{E}\Big[\sqrt{h_{t-1}}z_{t-1}\sqrt{h_t}\underbrace{\mathbb{E}[z_t|\mathcal{F}_t]}_{0}\Big] \\
&= 0.
\end{aligned}$$

Finally we have

$$\mathbb{E}[\sigma^2_{t+1,HR}] = a^2 + b^2\frac{\hat{\lambda}^2}{(1-e^{-\hat{\lambda}})^2}\sum_{i=1}^{\infty} h_{t-i}e^{-2i\hat{\lambda}}. \tag{2.28}$$

In comparison, the GARCH model gives us (equation (2.20) above),

$$\mathbb{E}[h_t] = \omega(1+\sum_{i=1}^{\infty}\beta^i) + \alpha\sum_{i=1}^{\infty}\beta^{i-1}h_{t-i}. \tag{2.29}$$

Note that the above GARCH model applies to a daily data and volatilities, whereas in the Hobson and Rogers model σ refers to an annualised volatility. In the numerical computation of the offset function we use end of day prices, and so set $k = 1/252$ (day^{-1}) and $\Delta = 1$ (day). In order to directly compare with equation (2.28) we need to apply a conversion factor to convert to annual volatilities. We define the annual volatility $\hat{h_t} = 252 \times h_t$, and so the above equation becomes

$$\mathbb{E}[\hat{h_t}] = 252\left(\omega(1+\sum_{i=1}^{\infty}\beta^i) + \alpha\sum_{i=1}^{\infty}\beta^{i-1}h_{t-i}\right). \tag{2.30}$$

We can then clearly see that

$$a^2 = 252\omega(1+\sum_{i=1}^{\infty}\beta^i). \tag{2.31}$$

Comparing the coefficients of h_{t-i} in the cases of $i = 1$ and $i = 2$ gives us

$$\hat{\lambda} = -\frac{\log(\beta)}{2}. \tag{2.32}$$

Finally we have

$$b^2 = \frac{252\alpha(1 - e^{-\hat{\lambda}})^2 e^{2\hat{\lambda}}}{\hat{\lambda}^2}. \tag{2.33}$$

2.3.1 Numerical Results

The data used was daily closing price data from the S&P500 from 3/2/2002 to 23/5/2003. GARCH(1,1) parameters were found using MATLAB and *garchfit()*. The calibration the GARCH(1,1) model gives

$$\begin{aligned} \omega &= 2.069 \times 10^{-5}, \\ \alpha &= 0.0766 \\ \beta &= 0.8316. \end{aligned}$$

Correspondingly we have, using equations (2.31), (2.32) and (2.33),

$$\begin{aligned} a &= 0.176 \\ \hat{\lambda} &= 0.0922 \\ \Rightarrow \lambda &= 23.23 \\ b &= 4.602 \end{aligned}$$

Now the Hobson and Rogers model is given by

$$\sigma^2_{HR} = a^2 + b^2 D_t^2.$$

Note that this is an annualised volatility. From the calibration of this model using a finite difference method in Chapter 3 we find for $\lambda = 23.23$

$$\begin{aligned} a &= 0.2214, \\ b &= 3.0027. \end{aligned}$$

Clearly the GARCH and Hobson and Rogers parameters do not coincide but are the same order of magnitude. It will be shown in Chapter 3 that the optimum choice of λ in the Hobson and Rogers model, for the above offset function, is given by $\lambda = 30$ which is also in approximate agreement with the GARCH comparison. The following table compares the coefficients of each model.

Table 2.1: Table of GARCH h_{t-k} coefficients

Coefficient	HR	GARCH
h_{t-1}	$\eta^2\gamma\left(\frac{\lambda}{1-e^{-\lambda}}\right)^2 e^{-2\lambda}$	α
h_{t-2}	$\eta^2\gamma\left(\frac{\lambda e}{1-e^{-\lambda}}\right)^2 e^{-4\lambda}$	$\beta\alpha + \alpha$
$\vdots$	$\vdots$	$\vdots$

Remark. *Discretisation of the second order offset function (setting $m = 2$) gives*

$$\sigma^2_{t+1,HR} = a^2 + b^2\left(\sum_{i=1}^{\infty}\hat{\lambda}e^{-\hat{\lambda}i}\left(\sum_{k=0}^{i}\epsilon_{t-k}\right)^2\right)^2.$$

2.4 Extensions of the model and Literature Review

We first discuss a discrete model by Jeantheau [23] in which the goal is to investigate the link between ARCH and the Hobson and Rogers model. Jeantheau uses a discrete version of the Hobson and Rogers model to motivate a new class of discrete models which are conditionally heteroscedastic and may be seen as an alternative to ARCH. Its diffusion approximation is shown to be a

complete stochastic volatility model. The diffusion approximation also provides a numerical scheme, different from the Euler scheme, to approximate the Hobson and Rogers model. Looking at the specification of volatility in the GARCH model, Jeantheau shows that we can write

$$\sigma_n^2 = \frac{\tilde{\omega}}{1-\tilde{\beta}} + \tilde{\alpha}\sum_{i\geq 1}\tilde{\beta}^{i-1}(\tilde{Z}_{n-i+1} - \tilde{Z}_{n-i})^2$$

Let us now describe Jeantheau's discrete model. Construct a process $Z_n, n \in \mathbb{Z}$, such that the offset functions satisfy

$$D_n^{(m)} = (1-\beta)\sum_{i=1}^{\infty}\beta^{i-1}(Z_n - Z_{n-i})^m.$$

The log of the price process now satisfies the discrete time version of (2.2), that is to say

$$Z_{n+1} - Z_n = \mu(D_n^{(1)}, \ldots, D_n^{(d)}) + \sigma(D_n^{(1)}, \ldots, D_n^{(d)})\eta_{n+1}$$

where $\eta_n, n \in \mathbb{Z}$, is defined as in the GARCH model. If $d = 2$ then it is shown that the discrete processes Z_t, $D_t^{(1)}$ and $D_t^{(2)}$ converge in distribution to the complete stochastic volatility model given by

$$\begin{aligned}
dZ_t &= \mu(D_t^{(1)}, D_t^{(2)})dt + \sigma(D_t^{(1)}, D_t^{(2)})dW_t \\
dD_t^{(1)} &= (\mu(D_t^{(1)}, D_t^{(2)}) - \lambda D_t^{(1)})dt + \sigma(D_t^{(1)}, D_t^{(2)})dW_t \\
dD_t^{(2)} &= (2D_t^{(1)}\mu(D_t^{(1)}, D_t^{(2)}) + \sigma^2(D_t^{(1)}, D_t^{(2)}) - \lambda D_t^{(2)})dt + 2D_t^{(1)}\sigma(D_t^{(1)}, D_t^{(2)})dW_t
\end{aligned}$$

Remark. *If we set $\beta = 1 - \lambda\Delta$, the second equation is the Euler scheme of the stochastic differential equation satisfied by the offset function of order 1. However, this is not the case for the offset function of order 2.*

Remark. *The assumption needed for convergence is also satisfied when*

$\beta_\Delta = exp(-\lambda\Delta)$, which is the choice inspired by the offset functions. In section 4 of [23] the case of

$$\sigma^2(D_n^{(2)}) = \omega + \alpha D_n^{(2)}$$

is considered. This results in the system of equations

$$\begin{aligned} D_{n+1}^{(1)} &= \beta D_n^{(1)} + (\omega + \alpha D_n^{(2)})^{1/2}\eta_{n+1}, \\ D_{n+1}^{(2)} &= \beta D_n^{(1)} + (\omega + \alpha D_n^{(2)})\eta_{n+1}^2 + 2\beta D_n^{(1)}(\omega + \alpha D_n^{(2)})^{1/2}\eta_{n+1} \end{aligned} \tag{2.34}$$

Jeantheau then goes on to prove this result and the following two propositions:

Proposition 18. *The system (2.34) admits a unique strictly stationary and positive recurrent solution with $E[D_n^{(2)}] < \infty$ if and only if $\alpha + \beta < 1$. In this case we have*

$$E[\sigma^2(D_n^{(2)})] = \frac{\omega(1-\beta)}{1-(\alpha+\beta)}$$

Proposition 19. *If $E[\ln(\beta + \alpha\eta_n^2)] < 0$, the system (2.34) admits a unique strictly stationary and positive recurrent solution. Moreover, there exists a $\delta \in (0, 1]$, such that $E[(D_n^{(2)})^\delta] < \infty$*

Finally we refer to the paper of Hubalek, Teichmann and Tompkins entitled 'Flexible Complete Models with Stochastic Volatility Generalising Hobson and Rogers' [24]. They investigate whether complete stochastic volatility models like the Hobson and Rogers model can produce appropriate smiles or not. Furthermore they suggest the following generalisation of the Hobson and Rogers model with remains complete but which, they claim, fits the features of actual

market data much better:

$$\begin{aligned} dZ_t &= -\frac{1}{2}\sigma_1(D_t)^2 dt + \sigma_1(D_t)dW_t, \\ dD_t &= \mu(D_t)dt + \sigma_2(D_t)dW_t, \\ Z_0 &= z, D_0 = d \end{aligned}$$

with the following specification

$$\begin{aligned} \sigma_1(d) &= \eta(1 + \epsilon\beta d^2) \\ \sigma_2(d) &= \xi\eta \\ \mu(s) &= -\frac{\eta^2}{2} - \lambda d \end{aligned}$$

for fixed $\epsilon > 0$. A solution of this generalised Hobson and Rogers model is then calculated.

Chapter 3

Calibration of the Hobson and Rogers model

In this chapter we study parameter estimation and numerical solutions of the Hobson and Rogers model. We revisit the work of Foschi and Pascucci [22] and examine how they numerically solve a transformed version of the Hobson and Rogers partial differential equation (2.16). The paper of Foschi and Pascucci provides a flexible calibration procedure in order to determine the parameters of the Hobson and Rogers model, specifically (λ, a, b) in the expression

$$\sigma^2(D_t) = a^2 + b^2 D_t^2.$$

where

$$D_t = D_t^{(1)} = Z_t - \int_0^\infty \lambda e^{-\lambda u} Z_{t-u} du$$

and Z_t is the log price. They use a finite-difference method to solve the partial differential equation (2.13) with initial condition (2.14) and then a non-linear least squares routine to vary the parameters of the model, to find those that give a best fit to market data. In this chapter we revisit the calibration

procedure used by Foschi and Pascucci for the above volatility specification. Then, in Section 3.6, we define new, more general volatility specifications not previously calibrated. One such specification is suggested in [25]. We carry out calibrations for a range of values of λ guided by our results from Chapter 2. We find that the choice $\lambda = 1$ as chosen in [22] is not necessarily optimal. We have found that the optimal values depend of the volatility specification being used and on the error metric specified in the calibration routine. This is discussed in more depth later.

Another aspect we investigate in this chapter is the time dependence of the volatility smile. Through extensive numerical investigations we have found that the implied volatility surface fluctuates daily. We conclude that the changing surface is a consequence of market sentiment. Since the volatility specification has no explicit time dependence it doesn't have the ability to reproduce this behaviour. In order to overcome this difficulty, we modify the dataset such that the average daily implied volatility is constant, and calibrate our model to this new dataset. The procedure is explained in detail in Section 3.7.3. The main results of this chapter are that using the optimal value of λ and adjusted data, we can provide an improved calibration routine where we see an order of magnitude reduction in residual errors.

The market data used, kindly provided by Paolo Foschi, is a set of S&P500 index option prices. This procedure can be carried out using the *lsqnonlin()* routine from MATLAB. This function requires the spatial derivatives of the option prices and market option prices, as well as the partial differential equation as inputs. It then provides the best fit parameters with errors. The primary MATLAB code used is very much based on code also provided by Paolo Foschi for which we are very grateful. We have used the finite difference scheme which was provided and the procedure for cleaning the data. We have reworked

the functions which calculate the value of the offset[1]. We found the method used in the non linear least squares routine to be unreliable. We found that speed of finding an optimum set of parameters is greatly increased by allowing MATLAB to determine the necessary Jacobian matrix in order to determine step size and direction when stepping through the parameter space. Details are provided in Section 3.6. Finally, modifications had to be made to take account of the extra parameters in the more general offset functions. Important sections of the code are provided in the Appendix. The full code is also available upon request.

3.1 The work of Foschi and Pascucci revisited

We start with the transformed equation (2.13) derived in the previous chapter, namely

$$\mathcal{L}V := a(V_{xx} - V_x) + (x - y)V_y - V_\tau = 0, \tag{3.1}$$

where $a = \sigma^2(D_t)/2\lambda$, and with boundary condition

$$V(x_T, y_T, 0) = (e^{x_T} - 1)^+. \tag{3.2}$$

The functional form of $\sigma(D_t)$ is specified below. The function $V(x, y, \tau)$ represents the price of an option under the following transformation

$$f(P_t, D_t, T - t) \to Ke^{r(T-t)}V(x, y, \tau), \tag{3.3}$$

[1]The MATLAB script written to compute the offset values is called *compute_trend()* and the corresponding code can be found in the Appendix.

where

$$\begin{aligned} x &= \log\left(\frac{P_t}{K}\right) - r(T-t), \\ y &= \log\left(\frac{P_t}{K}\right) - r(T-t) - D \text{ and} \\ \tau &= -\lambda(T-t). \end{aligned}$$

As usual, P_t is the price of the underlying, D_t is the Hobson-Rogers first offset function, $T-t$ is the time to maturity. The risk-free interest rate is r. The offset function is defined, as before, by

$$D_t^{(m)} = \int_0^\infty \lambda e^{-\lambda u}(Z_t - Z_{t-u})^m du,$$

where $Z_t = \log(e^{-rt}P_t)$ is the log-discounted price of the underlying. In the case of $m = 1$ we have

$$\begin{aligned} D_t = D_t^{(1)} &= \int_0^\infty \lambda e^{-\lambda u}(Z_t - Z_{t-u})du \\ &= \int_0^\infty \lambda e^{-\lambda u} Z_t du - \int_0^\infty \lambda e^{-\lambda u} Z_{t-u} du \\ &= Z_t - \int_0^\infty \lambda e^{-\lambda u} Z_{t-u} du. \end{aligned}$$

In their paper, Hobson and Rogers choose

$$\sigma^2(D_t) = a^2 + b^2 D_t^2.$$

We will review this choice and examine other choices. We follow and describe in detail here the finite-difference scheme provided in [22], using $m = 1$ in the offset function and pointing out any adjustments we make to the original implementation as we go.

3.2 Finite-Difference Operators

In the numerical approximation of the above p.d.e. the parts $\frac{\sigma^2(D_t)}{2\lambda}(V_{xx} - V_x)$ and $D_{\Delta u}V := (x-y)V_y - V_\tau$ are treated separately. We consider the uniform grid

$$G = \{(i\Delta_x, j\Delta_y, n\Delta_\tau) \quad | \quad i, j, n \in \mathbb{Z}, n \geq 0\}.$$

We use the following central difference approximation for ∂_x

$$\partial_x V(x,y,\tau) \cong D_{\Delta_x}V(x,y,\tau) = \frac{V(x+\Delta_x, y, \tau) - V(x-\Delta_x, y, \tau)}{2\Delta_x},$$

and three-point scheme for ∂_{xx}

$$\partial_{xx} V(x,y,\tau) \cong D^2_{\Delta_x}V(x,y,\tau) = \frac{V(x+\Delta_x, y, \tau) - 2V(x,y,\tau) + V(x-\Delta_x, y, \tau)}{\Delta_x^2}.$$

This leads to the approximation

$$\begin{aligned} \partial_{xx}V(x,y,\tau) - \partial_x V(x,y,\tau) &\cong D^2_{\Delta_x}V(x,y,\tau) - D_{\Delta_x}V(x,y,\tau) \\ &= d_1 V(x-\Delta_x, y, \tau) + d_2 V(x,y,\tau) + d_3 V(x-\Delta_x, y, \tau), \end{aligned} \tag{3.4}$$

with $d_1 = 1/\Delta_x^2 + 1/(2\Delta_x)$, $d_2 = -2/\Delta_x^2$ and $d_3 = 1/\Delta_x^2 - 1/(2\Delta_x)$.

The second main derivative is given by

$$\begin{aligned} D_{\vec{u}}V &= (x-y)V_y - V_\tau \\ &= \left(\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y}, \frac{\partial V}{\partial \tau}\right).(0, x-y, -1) \\ &= \nabla V.\vec{u}, \end{aligned}$$

where $\vec{u} = (0, x-y, -1)$ and with the directional derivative $D_{\vec{u}}V$ approximated

by either

$$\begin{aligned} Y_{\vec{u}}^{+}V &= \frac{\tilde{V}\left(x,y,\tau\right)-\tilde{V}\left(\left(x,y,\tau\right)-\Delta_{\tau}(0,x-y,-1)\right)}{\sqrt{1+(x-y)^{2}}\Delta_{\tau}} \\ &= \frac{\tilde{V}(x,y,\tau)-\tilde{V}\Big(\left(x,y-\Delta_{\tau}(x-y),\tau+\Delta_{\tau}\right)\Big)}{\sqrt{1+(x-y)^{2}}\Delta_{\tau}} \end{aligned}$$

or

$$\begin{aligned} Y_{\vec{u}}^{-}V &= \frac{\tilde{V}\left(\left(x,y,\tau\right)+\Delta_{\tau}(0,x-y,-1)\right)-\tilde{V}\left(x,y,\tau\right)}{\sqrt{1+(x-y)^{2}}\Delta_{\tau}} \\ &= \frac{\tilde{V}\Big(\left(x,y+\Delta_{\tau}(x-y),\tau-\Delta_{\tau}\right)\Big)-\tilde{V}(x,y,\tau)}{\sqrt{1+(x-y)^{2}}\Delta_{\tau}} \end{aligned} \tag{3.5}$$

Note that we divide by the norm of the vector $(0,x-y,-1)\Delta_{\tau}$ as standard when calculating the directional derivative. An alternative approach is given iin [22]. Details of this approach can be found in [26]. In dealing with a fixed uniform grid, the coordinates (x,y,t) and $(x,y+\Delta_{\tau}(x-y),t-\Delta_{\tau})$ may not both be grid points. To fix this problem we use linear interpolation between grid points in the y-direction. In the above approximations

$$\tilde{V}(x,y,\tau)=(1-\gamma)V(x,\tilde{y},\tau)+\gamma V(x,\tilde{y}+\Delta_{y},\tau),$$

where $\tilde{y}=[y/\Delta_{y}]\Delta_{y}$, with $\Big[\,\cdot\,\Big]$ denoting the integer part, and $\gamma=(y-\tilde{y})/\Delta_{y}$. Also

$$\begin{aligned} \tilde{V}(x,y,\tau) &= (1-\gamma)V(x,\tilde{y},\tau)+\gamma V(x,\tilde{y}+\Delta_{y},\tau) \\ &= (1-\gamma)V(x,y,\tau)+\gamma V(x,y+\Delta_{y},\tau) \qquad (\text{since } \tilde{y}=[y/\Delta_{y}]\Delta_{y}=y), \\ &= V(x,y,\tau) \qquad (\text{since } \gamma=(y-\tilde{y})/\Delta_{y}=0). \end{aligned}$$

The discrete operators $\mathcal{L}_G^+$ and $\mathcal{L}_G^-$ are defined by

$$\mathcal{L}_G^{\pm} = a\Big(D_{\Delta_x^2}V - D_{\Delta_x}V\Big) + Y_{\vec{u}}^{\pm}V.$$

$\mathcal{L}_G^+$ and $\mathcal{L}_G^-$ are the explicit and implicit schemes for the discretisation of the operator $\mathcal{L}$ in (3.1). We will work through the details of the finite difference scheme in a later section. Firstly we will look at the boundary conditions.

3.3 Boundary Conditions

The finite difference scheme discretises the system in a bounded region. We define the region in the cylinder

$$Q = \{(x, y, \tau) \ : \ |x| < \mu, |y| < \nu \quad \text{and} \quad \text{-}\lambda\text{T}{<}\tau{<}0\},$$

for some suitable large μ, ν. This corresponds to the initial-boundary value problem in the domain

$$\{(P, D, t) \ : \ |P| < Ke^{\mu+r(T-t)}, |D| < \nu, \text{and}\, 0 < t < T\}.$$

The conditions on the boundary of Q, defined by

$$\partial_P Q = \partial Q \cap \{(x, y, \tau) \mid -\lambda T < \tau < 0\},$$

are set as follows:

$$V(x, y, 0) = (e^x - 1)^+, \text{ for } x \in [-\mu, \mu],\, y \in [-\nu, \nu].$$

We set

$$(\partial_{xx}V - \partial_x V)(\pm\mu, y, \tau) = 0, \text{ for } y \in (-\nu, \nu),\ \tau \in (-\lambda T, 0). \tag{3.6}$$

This last condition is needed as the x boundary and the numerical approximation on this boundary requires points from outside the region. Effectively we are setting
$d_1 V(x - \Delta_x, y, \tau) = -\Big(d_2 V(x, y, \tau) + d_3 V(x + \Delta_x, y, \tau)\Big)$ at the lower boundary and $d_3 V(x + \Delta_x, y, \tau) = -\Big(d_1 V(x - \Delta_x, y, \tau) + d_2 V(x, y, \tau)\Big)$ at the upper boundary. This is a common approximation. Let us now introduce some shorthand notation. Fix $i_0,\ j_0 \in \mathbb{N} \cup \{0\}$, we denote

$$V^n_{i,j} = V(i\Delta_x, j\Delta_y, n\Delta_\tau), \quad i,\ j \in \mathbb{Z}, \quad |i| \le i_0, \quad |j| \le j_0.$$

Applying the notation to the operator (3.4) for $|i| \le i_0 - 1$ gives

$$D_{\Delta_x^2} V^n_{i,j} - D_{\Delta_x} V^n_{i,j} = (d_1 V^n_{i-1,j} + d_2 V^n_{i,j} + d_3 V^n_{i+1,j}).$$

Similarly we apply the notation to the operator in (3.5). We have, for $(x, y, \tau) = (i\Delta_x, j\Delta_y, n\Delta_\tau)$,

$$\tilde{V}\Big(x, y + \Delta_\tau(x + y), \tau - \Delta_\tau\Big) = (1 - \gamma)V^{n-1}_{i,j+k} + \gamma V^{n-1}_{i,j+k+1}$$

where

$$k = \left\lfloor \frac{(x - y)\Delta_\tau}{\Delta_y} \right\rfloor = \left\lfloor \Big(i\frac{\Delta_x}{\Delta_y} - j\Big)\Delta_\tau \right\rfloor \quad \text{and} \quad \gamma = \frac{(x - y)\Delta_\tau}{\Delta_y} - k \tag{3.7}$$

are the lower integer part and fractional part of $(x - y)\Delta_\tau/\Delta_y$ respectively. Now applying the $Y^-_{\Delta_\tau}$ operator to $V^n_{i,j}$ leads to

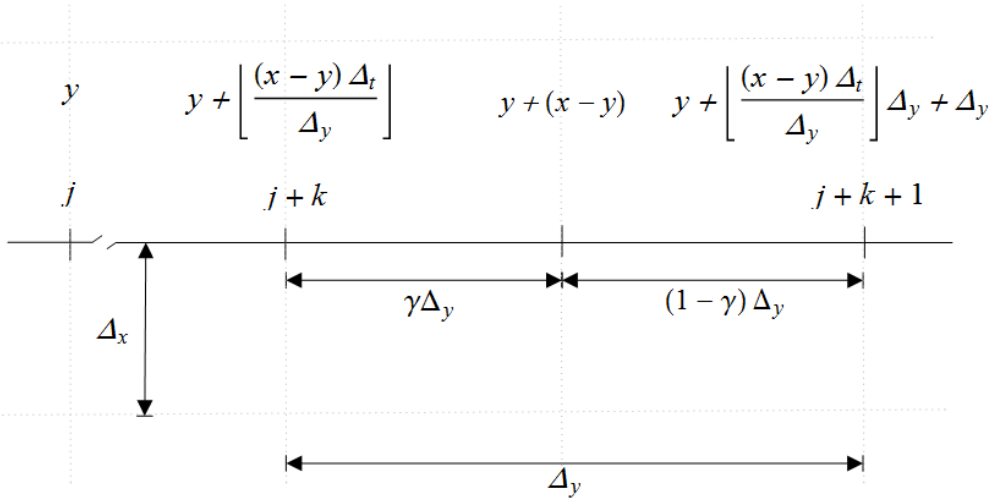

Figure 3.1: Setup of finite difference grid.

$$Y^-_{\Delta_\tau} V^n_{i,j} = \frac{1}{\Delta_\tau}\left((1-\gamma)V^{n-1}_{i,j+k} + \gamma V^{n-1}_{i,j+k+1} - V^n_{i,j}\right). \tag{3.8}$$

In Section 3.4 we show that with an appropriate choice of grid size, no boundary conditions need to be imposed in the y direction. This is rigorously proved in [22]. Applying the discrete operator $\mathcal{L}^-_G$ to $V^n_{i,j}$ reads

$$a_{i,j}\left(D_{\Delta_x^2} V^n_{i,j} - D_{\Delta_x} V^n_{i,j}\right) + Y^-_{\Delta_\tau} V^n_{i,j} = 0, \qquad |i| \le i_0 - 1,\ |j| \le j_0, \tag{3.9}$$

where $a_{i,j} = a(i\Delta_x, j\Delta_y)$. The condition (3.6) is equivalent to

$$Y^-_{\Delta_\tau} V^n_{i,j} = 0, \qquad i = \pm i_0, |j| \le j_0. \tag{3.10}$$

3.4 Numerical Scheme

With the numerical operators now defined, we work through the finite difference procedure. The discretisation of equation (2.13) is formulated here as a block diagonal linear system. We define $I = 2i_0 + 1$, $J = 2j_0 + 1$ and denote by $V^n \in \mathbb{R}^{IJ}$ the column vector containing the values $V^n_{i,j}$ for $|i| \le i_0$ and $|j| \le j_0$.

The values in the vector are sorted by their index pairs (j, i) in lexicographic order. Now consider the application of the discrete operator $Y^-_{\Delta_\tau}$ to the vector V^n. Note that $Y^-_{\Delta_\tau} V^n_{i,j}$ is a linear combination of the corresponding element in V^n and two elements, $V^{n-1}_{i,j+k}$ and $V^{n-1}_{i,j+k+1}$ of V^{n-1}. Thus, applying $Y^-_{\Delta_\tau}$ to V^n is equivalent to the difference of two linear operators, $\Delta_\tau^{-1}\mathcal{I}$ and $\Delta_\tau^{-1}Z$ applied respectively to V^n and V^{n-1}, where $\mathcal{I}$ denotes the identity operator in $\mathbb{R}^{IJ}$. Specifically $Y^-_{\Delta_\tau} V^n_{i,j}$ is given by

$$-\frac{1}{\Delta_\tau}(V^n - ZV^{n-1}), \tag{3.11}$$

where $Z \in \mathbb{R}^{IJ\times IJ}$ such that the entry corresponding to the index i, j of ZV^{n-1} is given by

$$(1-\gamma)V^{n-1}_{i,j+k} + \gamma V^{n-1}_{i,j+k+1}. \tag{3.12}$$

The selection of i_0 and j_0 is a numerical choice which specifies indirectly the number of grid points. We choose Δ_x such that $i_0\Delta_x = \max\Big(\log(\frac{S}{K}) - r(T-t)\Big)$. In our MATLAB script the notation is slightly different with the index running from 0 to $2\times(i_0+1)$ instead as this is easier to implement. At the edge of our grid the boundary condition (3.10) allows us to define the exterior values $V^n_{-i_0-1,j}$ and $V^n_{i_0+1,j}$ in terms of interior points. We may also choose j_0 such that $j_0\Delta_y > \max\Big(\log(\frac{S}{K}) - r(T-t) - D\Big)$. For a given $V^n_{i,j}$ the maximum j index upon which $V^n_{i,j}$ depends is $j+k+1$ at the time $n-1$, (see (3.8)). Let us look at how this subscript behaves. For $\Delta_\tau < 1$ the quantity $j+k+1$ is increasing linearly in j and linearly in i since

$$\begin{aligned} j+k+1 &= j + \left\lfloor \frac{(x-y)\Delta_\tau}{\Delta_y} \right\rfloor + 1 \\ &= j + \left\lfloor \frac{(i\Delta_x - j\Delta_y)\Delta_\tau}{\Delta_y} \right\rfloor + 1 \\ &= j + \left\lfloor \frac{i\Delta_x\Delta_\tau}{\Delta_y} - j\Delta_\tau \right\rfloor + 1. \end{aligned}$$

Given a grid of interest on which we want perform the finite difference scheme, we need to know how big this grid needs to be to include all dependencies $V_{i,j}^n$ which may not lie inside the original grid. The original grid is the grid defined by the full range of our market data. As described above, at the edges defined by $\pm i_0\Delta_x$ our finite difference scheme is defined such that the numerical derivatives do not depend on exterior points. Here we examine what happens at the boundary $\pm j_0\Delta_y$. As we move forward in time in our numerical scheme, the point $V_{i,j}^n$ will depend upon $V_{i,j+k+1}^{n-1}$ and $V_{i,j+k}^{n-1}$. Looking at how k varies across the grid, its minimum value can be expressed as

$$\min(k) = (-x_0 - j_0\Delta_y)\Delta_\tau/\Delta_y \tag{3.13}$$

A numerical check shows that for a grid of 101×81 and with the given market data, that $\min(k) \sim -2.88$, and this occurs on the $+j_0$ boundary. Similarly we have $\max(k) \sim 0.02$ and this occurs on the $-j_0$ boundary. In the case of k falling between integers, the interpolation scheme described above is used and so the minimum dependence at the $+j_0$ boundary is $j-3$ while at the $-j_0$ boundary it is $j+1$. Both of these results show that $V_{i,j}^n$ can be computed from points entirely within the grid at time $n-1$. This further shows that there is no need for boundary conditions at $\pm j_0$. We now refer back to equation (3.9) and, following Foschi and Pascucci, we use algebraic manipulation to rewrite the system as one that is computationally easier to solve. We have

$$\begin{aligned} a_{i,j}\Big(D_{\Delta_x^2}V_{i,j}^n - D_{\Delta_x}V_{i,j}^n\Big) + Y_{\Delta_\tau}^- V_{i,j}^n &= a_{i,j}\Big(D_{\Delta_x^2} - D_{\Delta_x}\Big)V_{i,j}^n + \frac{1}{\Delta_\tau}\Big(ZV^{n-1} - \mathcal{I}V^n\Big) \\ &= \Big[-\mathcal{I} + a_{i,j}\Delta_\tau(D_{\Delta_x^2} - D_{\Delta_x})\Big]V_{i,j}^n + ZV^{n-1} \\ &= 0. \end{aligned}$$

Let $A \in \mathbb{R}^{IJ\times IJ}$ be the diagonal matrix with elements $a_{i,j}$. Let the matrices

$$D = \begin{pmatrix} \breve{D} & 0 & \cdots & 0 \\ 0 & \breve{D} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \breve{D} \end{pmatrix} \qquad \text{and} \qquad \breve{D} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 \\ d_1 & d_2 & d_3 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & d_1 & d_2 & d_3 \\ 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

be tridiagonal matrices of order IJ and I respectively. Thus the matrix $D \cdot V^n$ represents the operator $(D_{\Delta_x^2} - D_{\Delta_x})V^n$ with the boundary condition (3.6) $(\partial_{xx}V - \partial_x V)(\pm\mu, y, \tau) = 0$ built in. We rewrite the operator $-\mathcal{I} + a_{i,j}\Delta_\tau(D_{\Delta_x^2} - a_{i,j}D_{\Delta_x})$ as $-\mathcal{I} + \Delta_t A \cdot D$, which can be rewritten as

$$\bar{A}_1 \cdot V^n = \bar{A}_2 \cdot V^{n-1}, \qquad 1 \leq n \leq N, \tag{3.14}$$

with $\bar{A}_1 = \{-\mathcal{I} + \Delta_t A \cdot D\}$ and $\bar{A}_2 = -Z_n$. We note here that the above expressions differ from those of Foschi and Pascucci in that firstly, we have a fixed grid size for all time-steps, thus no subscript n and, secondly, we define $\bar{A}_1$ and $\bar{A}_2$ with the sign change discussed in Section 3.3. For each n, the system (3.14) can be written as a linear system:

$$\begin{pmatrix} I & 0 & 0 & \cdots & 0 \\ -\bar{A}_2 & \bar{A}_1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \cdots & -\bar{A}_2 & \bar{A}_1 & 0 \\ 0 & \cdots & 0 & -\bar{A}_2 & \bar{A}_1 \end{pmatrix} \cdot \begin{pmatrix} V^0 \\ V^1 \\ \vdots \\ V^{N-1} \\ V^N \end{pmatrix} = \begin{pmatrix} V^i \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix},$$

or

$$\bar{A}\bar{V} = V^i$$

where V^i is the initial data i.e. $I \cdot V^0 = V^i$. Note that $n = 0$ corresponds to the maturity date of the option, and $n = N$ corresponds to the time of writing of the option.

Example 1.

In order to reinforce these ideas, let us work through a toy example. We consider three time steps, and a 3×3 spatial grid. The initial conditions are given by

$$V(x, y, 0) = (e^x - 1)^+.$$

This example is for demonstration purposes so the region in which the finite difference scheme is evaluated here is chosen for ease of demonstration and is not supposed to reflect values corresponding to the financial world. Let our x and y values range from 0 to 2, with $\Delta_x = 1, \Delta_y = 2$, and $\Delta_\tau = 0.5$. We set $i_0 = j_0 = 1$ such that our grid points lie on the integers running from $-1, 0, 1$. Correspondingly we have $I = J = 3$. At expiry we have $n = 0$ and our initial values (value at maturity) are given by the above equation. We have

$$V^0 = \begin{pmatrix} 0 & 0 & 0 \\ 1.72 & 1.72 & 1.72 \\ 6.38 & 6.38 & 6.38 \end{pmatrix}.$$

Then

$$\begin{aligned} V^1 &= \bar{A}_1^{-1} \cdot \bar{A}_2 \cdot V^0 \qquad \text{by (3.14)} \\ &= \bar{A}_1^{-1} \cdot (-Z) \cdot V^0 \end{aligned}$$

Let us first deal with the $Z \cdot V^0$ component. We need to compute the Z

operator which depends on our (x, y) values through k, see equation (3.12).

$$ZV^{n-1} = (1-\gamma)V^{n-1}_{i,j+k} + \gamma V^{n-1}_{i,j+k+1}.$$

Now $k = \left\lfloor \frac{(i\Delta_x - j\Delta_y)}{\Delta_y}\Delta_\tau \right\rfloor$ so we can determine $j+k$ for each point in our 3×3 grid. We have

$$\bar{k} = \begin{pmatrix} k_{-1,-1} & k_{-1,0} & k_{-1,1} \\ k_{0,-1} & k_{0,0} & k_{0,1} \\ k_{1,-1} & k_{1,0} & k_{1,1} \end{pmatrix} = \begin{pmatrix} 0 & -1 & -1 \\ 0 & 0 & -1 \\ 0 & 0 & -1 \end{pmatrix}.$$

and consequently $j+k, j+k+1$ values are given by

$$\begin{pmatrix} (-1,0) & (-1,0) & (0,1) \\ (-1,0) & (0,1) & (0,1) \\ (-1,0) & (0,1) & (0,1) \end{pmatrix}.$$

We construct the Z matrix using the information given in the above array.

$$\begin{pmatrix} 1-\gamma_{-1-1} & 0 & 0 & \gamma_{-1-1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 1-\gamma_{0-1} & 0 & 0 & \gamma_{0-1} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1-\gamma_{1-1} & 0 & 0 & \gamma_{1-1} & 0 & 0 & 0 \\ 1-\gamma_{-10} & 0 & 0 & \gamma_{-10} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-\gamma_{00} & 0 & 0 & \gamma_{00} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1-\gamma_{10} & 0 & 0 & \gamma_{10} \\ 0 & 0 & 0 & 1-\gamma_{-11} & 0 & 0 & \gamma_{-11} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-\gamma_{01} & 0 & 0 & \gamma_{01} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1-\gamma_{11} & 0 & 0 & \gamma_{11} \end{pmatrix}.$$

We must now determine γ for the above matrix, which is given by equation (3.7)

$$\gamma = \frac{(x-y)\Delta_\tau}{\Delta_y} - k,$$

and so is a function of (i,j). We calculate γ for each (i,j) and store it in the matrix $\bar{\gamma}$.

$$\bar{\gamma} = \begin{pmatrix} 0.25 & 0.75 & 0.25 \\ 0.5 & 0 & 0.5 \\ 0.75 & 0.25 & 0.75 \end{pmatrix}.$$

Finally we have

$$\bar{A}_2 \cdot V^0 = \begin{pmatrix} -V^0_{-1-1}(1-\gamma_{-1-1}) - V^0_{-10}\gamma_{-1-1} \\ -V^0_{0-1}(1-\gamma_{0-1}) - V^0_{00}\gamma_{0-1} \\ -V^0_{1-1}(1-\gamma_{1-1}) - V^0_{10}\gamma_{1-1} \\ -V^0_{-1-1}(1-\gamma_{-10}) - V^0_{-10}\gamma_{-10} \\ -V^0_{00}(1-\gamma_{00}) - V^0_{01}\gamma_{00} \\ -V^0_{10}(1-\gamma_{10}) - V^0_{11}\gamma_{10} \\ -V^0_{-10}(1-\gamma_{-11}) - V^0_{-11}\gamma_{-11} \\ -V^0_{00}(1-\gamma_{01}) - V^0_{01}\gamma_{01} \\ -V^0_{10}(1-\gamma_{11}) - V^0_{11}\gamma_{11} \end{pmatrix} = - \begin{pmatrix} 0 \\ 1.71828 \\ 6.38906 \\ 0 \\ 1.71828 \\ 6.38906 \\ 0 \\ 1.71828 \\ 6.38906 \end{pmatrix}.$$

Next we determine $(\bar{A}_1)^{-1}$. We have from before

$$\bar{A}_1 = \{-\mathcal{I}^9 + \Delta_t A \cdot D\}$$

which is a 9×9 matrix whose diagonal elements are

$$diag = \begin{pmatrix} -1 \\ (0.5a_{21}d_2 - 1) \\ -1 \\ -1 \\ (0.5a_{22}d_2 - 1) \\ -1 \\ -1 \\ (0.5a_{32}d_2 - 1) \\ -1 \end{pmatrix},$$

and upper diagonal given by

$$upper = \begin{pmatrix} 0 \\ (0.5a_{21}d_3) \\ 0 \\ 0 \\ (0.5a_{22}d_3) \\ 0 \\ 0 \\ (0.5a_{32}d_3) \end{pmatrix}$$

and lower diagonal given by

$$lower = \begin{pmatrix} (0.5a_{21}d_1) \\ 0 \\ 0 \\ (0.5a_{22}d_1) \\ 0 \\ 0 \\ (0.5a_{32}d_1) \\ 0 \end{pmatrix}.$$

The inverse of the above matrix, $(\bar{A}_1)^{-1}$, is too cumbersome to be shown here but is easily determined using software. For simplicity, setting $a_{ij} = 1$ for all (i, j). We find

$$\bar{A}_1 = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.75 & -2. & 0.25 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.75 & -2. & 0.25 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.75 & -2. & 0.25 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix},$$

and finally

$$V^1 = \bar{A}_1^{-1} \cdot \bar{A}_2 \cdot V^0 = \begin{pmatrix} V^1_{1,1} \\ V^1_{2,1} \\ V^1_{3,1} \\ V^1_{1,2} \\ V^1_{2,2} \\ V^1_{3,2} \\ V^1_{1,3} \\ V^1_{2,3} \\ V^1_{3,3} \end{pmatrix} = \begin{pmatrix} 0 \\ 1.65777 \\ 6.38906 \\ 0 \\ 1.65777 \\ 6.38906 \\ 0 \\ 1.65777 \\ 6.38906 \end{pmatrix}$$

To determine V^n we iteratively apply the same operators (which do not need to be recomputed at each iteration) :

$$V^n = \bar{A}_1^{-1} \cdot \bar{A}_2 \cdot V^{n-1} \tag{3.15}$$

3.5 Implementation of the calibration procedure.

The calibration of the Hobson and Rogers model is the estimation of the volatility function σ from observed market prices of European options. We assume $\sigma = \sigma(\cdot;\alpha)$ depends on a vector $\alpha = (\alpha_1, \ldots, \alpha_p)$ of real positive parameters and denote by $V(x, y, t; \alpha)$ the solution to the problem (3.1)-(3.2) corresponding to

$$a = \frac{\sigma^2(x-y;\alpha)}{2\lambda}.$$

Let $\hat{f}_i$ be the observed option value at the point $z_i = (x_i, y_i, t_i)$, for $i = 1, 2, \ldots, M$, and let $f_i(\alpha)$ be the price given by (3.3) in terms of the solution $V(x_i, y_i, \tau_i; \alpha)$ of the p.d.e. (3.1) for a given α at the observation point z_i. If the point z_i does not belong

to the grid G, the value of $f_i(\alpha)$ is approximated by linear interpolation between the nearest points on the grid in the $y-$direction.

The objective of the calibration is to minimise some measure of error of the model by choice of α. We have considered a number of error metrics. These are 'Root Mean Square Error'(RMSE), 'Mean Absolute Error'(MAE), 'Root Mean Square Relative Error'(RMSRE), and 'Mean Relative Absolute Error'(MRAE). These are defined as follows,

$$\begin{aligned}
\text{RMSE} &= \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f_i - \hat{f}_i)^2}, \\
\text{MAE} &= \frac{1}{N}\sum_{i=1}^{N}|f_i - \hat{f}_i|, \\
\text{RMSRE} &= \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{f_i - \hat{f}_i}{\hat{f}_i}\right)^2} \quad \text{and,} \\
\text{MRAE} &= \frac{1}{N}\sum_{i=1}^{N}\frac{|f_i - \hat{f}_i|}{\hat{f}_i}
\end{aligned}$$

We also consider the error metric

$$\text{RMSE}^{\text{original}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(v_i - \hat{v}_i)^2},$$

where v and $\hat{v}$ corresponds to the option prices in the transformed space. This appears to be the error metric used by Foschi and Pascucci in their calibration algorithm. For a chosen error metric for the calibration, we determine the optimal parameter set α. As a final step, we calculated the residual error as defined be each of the above metrics for the optimal α. Another point to note is that the parameter space over which we calibrate also includes λ. We find that given a volatility specification, and an choice of error metric, that different values of λ are optimal. Typically we find that $0 < \lambda < 10$. Details follow in a later section.

We will now go through in detail how we implement the calibration procedure in MATLAB. There are two main parts to the overall algorithm. The first is an algorithm to solve the transformed equation (3.1) subject to the boundary condition

(3.2) and the second is the algorithm to apply the non-linear least-squares routine.

3.5.1 The Dataset and MATLAB code.

The dataset was kindly provided by Paolo Foschi. The data is a set of European option quotations on the S&P 500 index from the Chicago Board Options Exchange. We now examine in detail the MATLAB algorithm used in [22] to calibrate the p.d.e. to the data. An outline of the calibration algorithm can be found in the Appendix. The first step is to convert the raw data into a usable format. The objective here is to have each option price associated with its corresponding underlying price, offset, strike, and time to maturity. In determining the offset we must choose a value of λ. In Section 3.6.1 we will discuss choices of λ and suggest how to determine the 'best' value of λ for a given dataset. Recall that the offset function is given by

$$D_t = D_t^{(1)} = \int_0^\infty \lambda e^{-\lambda u}(Z_t - Z_{t-u})du.$$

Unless otherwise stated, the MATLAB functions referred to in the following text are partly new work and their code can be found in the Appendix. The function names are the same as those MATLAB functions used by Foschi and Pascucci which allows for easier integration with Foschi and Pascucci's finite difference scheme.

The *compute_trend()* function is used to compute the offset at each time. It outputs a column vector where the ith row gives the value of the offset function at time i. The next step of the algorithm is to apply the transformations of equation (3.3) to our data. Recall that our new variables are defined as follows

$$f(P, D, T-t) \quad \rightarrow \quad Ke^{r(T-t)}V(x, y, \tau)$$

where

$$\begin{aligned} x &= \log(\frac{P}{K}) - r(T-t) \\ y &= \log(\frac{P}{K}) - r(T-t) - D, \qquad \text{and} \\ \tau &= -\lambda(T-t). \end{aligned}$$

The choice of λ is discussed in Section 3.6.1. Our raw data is now correctly parameterised and we use the function *calibrate()*[2] to choose $\alpha = (\alpha_1, \ldots, \alpha_p)$ such that the solution of the p.d.e. (2.13) is as close as possible to the market data. We use the built-in MATLAB function *lsqnonlin()* which attempts to minimise the squared sum of residuals by varying the vector α, as described in the algorithm above. We perform a number of calibrations, looking at different specifications of the volatility function, including that as originally specified in [1].

3.6 Choice of volatility function

A key question is to decide upon the functional form of the volatility function. Naturally we start with that proposed by Hobson and Rogers in their original paper [1]. It is specified by

$$\sigma_{HR}^2 = \min\left\{\alpha_1 + \alpha_2(D_t - \alpha_3)^2, \sqrt{5}\right\}.$$

This is the function also used in the calibration paper of Foschi and Pascucci [22]. Note that the above formulation is slightly more general than that analysed by Hobson and Rogers, who do not include a linear term, corresponding to 'Vol_3' below. We label the above volatility specification as 'Vol_0'. We also work with a number of

[2]The structure of this function remains mostly unchanged from that provided by P. Foschi. Alterations are made where we have modified the use of the nonlinear least squares function (internal MATLAB function) and where we have adapted the function to allow the use of more general volatility specifications.

other functional forms, namely

$$\begin{aligned} Vol_1 : \sigma^2 &= \frac{\alpha_1 + \alpha_2 D_t^2}{\alpha_4 + \alpha_5 D_t^2}, \\ Vol_2 : \sigma^2 &= \frac{\alpha_1 + \alpha_2 (D_t - \alpha_3)^2}{\alpha_4 + \alpha_5 (D_t - \alpha_6)^2}, \\ Vol_3 : \sigma^2 &= \alpha_1 + \alpha_2 D_t^2 \end{aligned}$$

The indices on the α coefficients are chosen for ease of comparison post calibration i.e. α_2 will always correspond to the coefficient of D_t^2 in the numerator. The function *evaluate()*[3] was written to compute the residual vector used in the calculation of each of the error metrics defined above.

Foschi and Pascucci in [22] provide a p.d.e., the solution of which gives the Jacobian of the option price values with respect to the parameters in the volatility specification. This is then used by the *lsqnonlin()* function to determine step size and direction which searching the parameter space for the optimal parameter values. This method has benefits in that it allows quicker computation of the Jacobian. We have found, though, that convergence is quicker and more reliable if MATLAB is allowed to compute the Jacobian using it own internal functionality. The slower computation is compensated by the quicker convergence.

The *evaluate()* function uses the functions *HR_calibrate()*[4] and *kolmogorov()*[5] to compute the solution to the p.d.e. on the grid, subject to initial conditions and boundary conditions. These functions are also given in the appendix.

3.6.1 Choice of λ

The one parameter that isn't determined by the calibration process is λ, which determines the rate of decay of the discount factors found in the offset function. We have previously attempted to quantify λ by making a direct comparison between the

[3] Adapted from work of P. Foschi
[4] Adapted from work of P. Foschi
[5] Provided by P. Foschi

offset function and a GARCH model. One approach we have decided to investigate is to see which λ minimises each of the error metrics, for each of the volatility functions defined above.

We found that there is an optimal choice of λ in the case of volatility specifications Vol_0 and Vol_2. This optimal choice of λ is given in each of the results tables. We perform a joint calibration procedure, recalibrating the model for a range of λ values. We define optimal in the sense that this choice of λ minimises the chosen error metric. A discussion on this can be found in Chapter 2, Section 3 in which the GARCH model suggests we that the optimal choice is given by $\lambda \simeq 23$.

In the case of Vol_1 and Vol_3, optimal lambda values occurred for $0.001 < \lambda < 0.1$. In both these cases the residual errors from the calibration were large when compared with those found by choosing either Vol_0 or Vol_2. For example, the $\text{RMSE}^{\text{original}}$ for the optimal λ in the case of Vol_1 was found to be of the order of $\simeq 7 \times 10^{-3}$ while in the case of Vol_0 or Vol_2 the same metric was of the order of $\simeq 4.4 \times 10^{-3}$ for an optimal choice of λ. We found this as a result of extensive numerical testing. It is notable that in both these specifications we only have a D_t^2 term. By squaring the offset we lose information about the direction of price movements. The extra D_t term in Vol_0 and Vol_2 provides this information. These results have led us to drop Vol_1 and Vol_3 from our investigations from here on. Notably, by doing this, we are eliminating the specification suggested originally by Hobson and Rogers in [1].

An interpretation of the results can be found in Section 3.7.4.

Note that in [22] the authors take $\lambda = 1$. In a later section we will describe a calibration method using 'adjusted data'. In that case the optimal choice of λ is given by $1 < \lambda < 5$. See Figure 3.3.

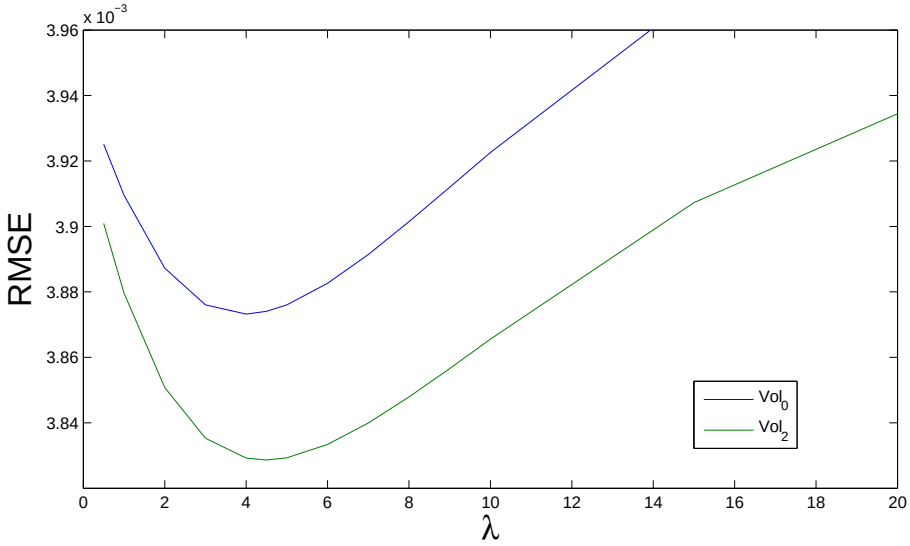

Figure 3.2: Original data set: A plot of $\text{RMSE}^{\text{original}}$ from the calibration routine vs. λ. We find that for Vol_0 and Vol_2 the optimal choice is given by $\lambda \simeq 4$.

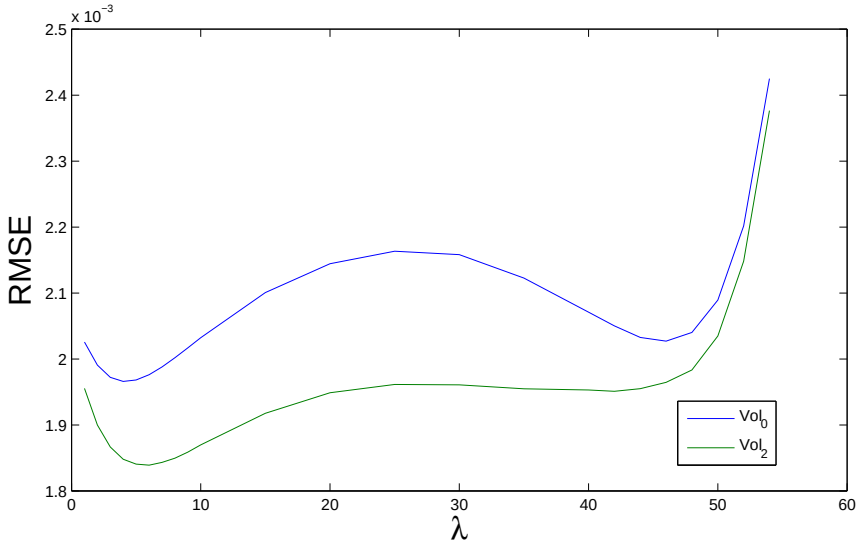

Figure 3.3: Adjusted data set: A plot of $\text{RMSE}^{\text{original}}$ from the calibration routine vs. λ. We find that for Vol_0 and Vol_2 the minimum occurs at $\lambda \simeq 4$ and $\lambda \simeq 5$ respectively.

3.7 Calibration Results

The overall aim of our model is to incorporate the information we have about the value of the offset function, in order to provide more accurate option valuations. It is common practice in the financial industry to trade options based on their Black-Scholes implied volatility as opposed to actual option price. As such, the

calibration of stochastic volatility models is typically implemented such that the implied volatility smile produced by the model matches the implied volatility smile of the market and this also forms the basis for testing the validity of a given model. This process is known as 'roundtripping'. In our calibration routine we try to minimise over the set of error metrics defined earlier, as opposed to implied volatility. We present the results of the calibration of the parameters in the volatility specifications Vol_0 and Vol_2. We then convert model and market prices into implied volatilities and compare the resulting data sets. This section also contains some statistics relating to these comparisons.

3.7.1 Initial Results

We now present the values found by the above calibration procedure for each of the above volatility specifications calibrated to each of the metrics defined above.

Table 3.1: Best fit parameters for the Vol_0 volatility specification for each of the error specifications.

Vol_0	$RMSE^{original}$	$RMSRE(\%)$	$RMSE$	MAE	$MRAE$
λ	4	1	3	4	1
α_1	$0.04 \pm 4.5E-07$	$0.04 \pm 2.0E-07$	$0.04 \pm 4.0E-07$	$0.04 \pm 1.6E-07$	$0.04 \pm 7.3E-08$
α_2	$4.94 \pm 2.6E-04$	$2.52 \pm 4.3E-04$	$4.1 \pm 2.3E-04$	$4.52 \pm 6.3E-05$	$2.53 \pm 6.4E-05$
α_3	$0.07 \pm 1.1E-03$	$0.09 \pm 1.3E-03$	$0.08 \pm 1.2E-03$	$0.08 \pm 6.6E-04$	$0.11 \pm 7.8E-04$

Table 3.2: Best fit parameters for the Vol_2 volatility specification for each of the error specifications.

Vol_2	$RMSE^{original}$	$RMSRE(\%)$	$RMSE$	MAE	$MRAE$
λ	4	1	4	4	2
α_1	$130.82 \pm 1.8E06$	$150.3 \pm 7.4E07$	$148.87 \pm 7.7E06$	$258.55 \pm 1.8E-01$	$361.7 \pm 2.8E-03$
α_2	$2542.13 \pm 3.5E07$	$4606.31 \pm 2.3E09$	$3617.91 \pm 1.9E08$	$3133.11 \pm 8.3E-01$	$7186.31 \pm 6.0E-01$
α_3	$0.11 \pm 1.2E-02$	$0.05 \pm 1.4E-02$	$0.09 \pm 7.5E-03$	$0.22 \pm 3.4E-03$	$0.09 \pm 7.9E-04$
α_4	$9525.34 \pm 1.3E08$	$7451.84 \pm 3.7E09$	$9902.49 \pm 5.1E08$	$24248.12 \pm 2.8E01$	$19204.14 \pm 4.0E00$
α_5	$-0.5 \pm 1.2E-01$	$-0.6 \pm 4.4E-01$	$-0.52 \pm 1.9E-01$	$-0.48 \pm 4.3E-03$	$-0.56 \pm 7.9E-03$

Table 3.3: Residual error values for each of the error metrics. The volatility specification here is Vol_0. Each column denotes a specific metric, while each row presents the residual error under each of those metrics.

Vol_0	$RMSE^{original}$	$RMSRE(\%)$	$RMSE$	MAE	$MRAE$
$RMSE^{original}$	0.0044	0.0047	0.0044	0.0045	0.0044
$RMSRE(\%)$	13.11%	12.67%	12.96%	13.55%	13.2%
$RMSE$	13.99	15.4	13.94	14.61	14.26
MAE	\$2.78	\$3.15	\$2.81	\$2.75	\$2.82
$MRAE$	.071	.0742	.0709	.0709	.0704

Table 3.4: Residual error values for each of the error metrics. The volatility specification here is Vol_2.

Vol_2	$RMSE^{original}$	$RMSRE(\%)$	$RMSE$	MAE	$MRAE$
$RMSE^{original}$	0.0044	0.0046	0.0044	0.0044	0.0044
$RMSRE(\%)$	13.19%	12.65%	13.04%	13.78%	13.18%
$RMSE$	13.81	14.84	13.77	14.42	13.99
MAE	\$2.75	\$2.96	\$2.77	\$2.72	\$2.76
$MRAE$	.0703	.0731	.0705	.0705	.07

For comparison, in [22] it was found that the optimal parameter choice and resulting residuals are given in the Table 3.5.

Before analysing the results there are a number of differences to note in the calibration procedures and in how the results are presented. In the MATLAB code that underlies the calibration results from [22] it seems that the calibration was implemented such that the residual being minimised was a root mean square error of

Table 3.5: Calibration results of Foschi and Pascucci. Note that the parameter λ pre calibration. These parameters may be compared with those in column 1 of Table 3.1 and the diagonal elements of Table 3.3

Vol_0			
RMSE	\$1.857	λ^*	1
MAE	\$1.463	α_1	$0.0272 \pm 8.51 \times 10^{-5}$
RMSRE	5.19%	α_2	$0.7114 \pm 2.11 \times 10^{-3}$
MRAE	3.34%	α_3	0.0616 ± 10^{-4}

the option values in the transformed space. We note that this approach may not give the same results as the case where the residuals are defined in terms of the untransformed variables. In the case of the relative errors i.e. RMSRE and MRAE, the approaches are equivalent due to cancellation of the conversion factors between the transformed and untransformed variables.

The second point to note is that only a single the error metric is used in the calibration algorithm of Foschi and Pascucci [22]. In contrast, we calibrate a new set of parameters, including λ, for each error metric in the untransformed variable space, and using the untransformed model and market option prices, we calculate RMSE, MAE, RMSRE and MRAE. It is clear for instance that the choice of error metric used in the calibration will affect the model prices and the value of each of the four measures of error. Choosing RMSRE as the error metric in the calibration for example, will mean the calculated RMSRE post calibration, will be minimised with respect to RMSRE calculated via other optimization error metrics. It is possible to see this effect in Tables 3.3 to 3.11 by looking at the diagonal entries in each of these arrays and noting that the diagonal element is less than or equal to any other entry in a given row. In some cases the differences across a given row are significant. It is up to the user of the model to decide which error metric to choose in the calibration.

Another point to note is that in calculation of the relative errors, [22] remove options from the market data set with value less than \$10. This is to reduce percentage bias which would appear for smaller option values. Indeed an examination of the market data vs our own model prices confirmed this effect to be significant. In our calibration we exclude options below the \$10 barrier for all calibrations. We will come back to

this issue in Chapter 5, where we will also discuss the parameter ranges in which we found the model to be weak. The model is particularly weak for longer maturities, and deep out of the money options. This may be an artifact of bad or unclean data and these topics will be discussed.

In the results of [22] a calibration with respect to three alternative models is presented. These consist of the standard model as presented in Vol_0, with a single calibration, a Vol_0 model with daily recalibration, and a third model based on a spline interpolation scheme. The second model in this list may be comparable to the model in which we calibrate to adjusted data. As will be discussed later, the adjusted dataset is constructed such that daily variability is eliminated, allowing us to avoid daily recalibration. We describe how to convert back to unadjusted prices. Where adjusted data was used in our calibrations, the measures of fit used in the optimization routine were calculated with respect to unadjusted data.

Finally, we calibrate to mid-option prices and do not consider 'outside' errors, where the error is defined as the maximum distance between the bid and ask prices if the model price is outside the bid-ask range, and is take as zero otherwise. Clearly this will lead to a fit that appears to be better. It may be prudent when performing a calibration to calibrate to either the bid or ask depending on the makeup of your current portfolio. For example, if the model is to be used to value a portfolio, the value of the asset to you may be best represented by the cost of selling that asset if holding a long position, or buying that asset if holding a short position. The model should reflect the cost of liquidating the position and so should be calibrated to the appropriate bid or ask prices. This fine tuning may be more relevant to those who wish trade more frequently and is a modelling choice.

Now, analysing the results, we can see that our parameters compare well with the calibrated paymasters reported in [22] in the case of Vol_0. We note that α_1 and α_3 are particularly close, while α_2 is a factor of two or four larger depending on whether comparing with Vol_0 or Vol_2. The parameters are still of the same order of magnitude which is reassuring. When comparing these results note that we have

different values of λ associated with each. The λ in our model is the result of a further manual calibration as discussed in the previous section. We found $\lambda = 4$ to be optimal in this case whereas in [22] the authors take $\lambda = 1$.

A higher value of λ will reduce the value of the offset value process at any given time due to larger discounting within the offset function. Since α_2 is the parameter that scales the offset value in our model, a higher value of α_2 would compensate for a reduced offset value. A distribution of the errors between market and model prices can be found in Figure 3.4 and the corresponding statistics for Vol_0 and Vol_2 in Tables 3.6 and 3.7. It shows us that both these models overestimate the option implied volatilities by between 0.24% and 0.34%. Surprisingly here, while Vol_2 has a lower $\text{RMSE}^{original}$ when looking at option prices, when looking at implied volatilities Vol_0 fares better. We also note that if the error distribution is fitted to a normal distribution we see a standard deviation in both models of $\simeq 3.4\%$. This is something that is worrying and suggests that we need to find ways to improve the model.

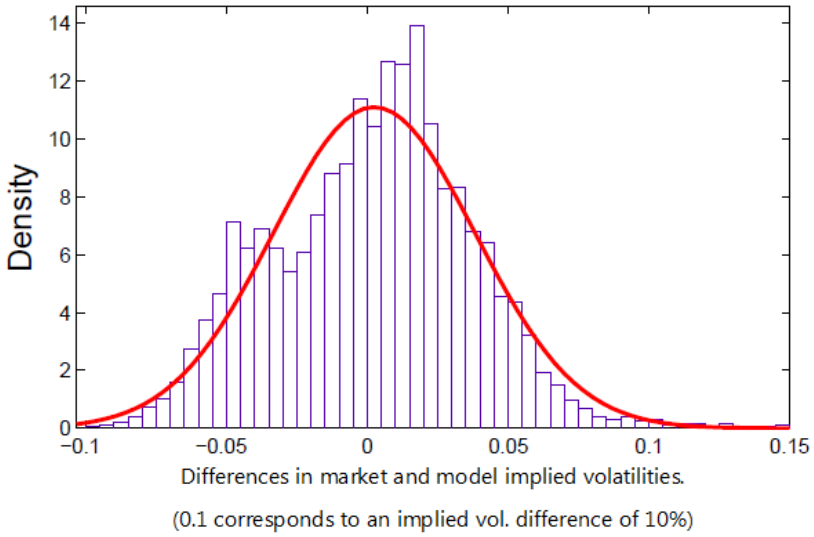

Figure 3.4: Distribution of differences between market implied volatility and model implied volatility in the case of Vol_0, with a fitted normal density function. See Table 3.6 for the mean and standard deviation of this fit.

In Figure 3.5 we graph market and model option prices as a function of moneyness where moneyness is defined as $\log(F_t/K)/\sqrt{T-t}$, where $F_t = e^{r(T-t)}P_t$. The moneyness variable is intended to allow us to compare option implied volatilities as a

Table 3.6: Parameters for normal distribution fitted to a distribution of errors between market and model implied volatilities in the case of Vol_0

Vol_0	Parameter Estimate	Std. Error
μ	0.24%	0.058%
σ	3.5%	0.041%

Table 3.7: Parameters for normal distribution fitted to a distribution of errors between market and model implied volatilities in the case of Vol_2

Vol_2	Parameter Estimate	Std. Error
μ	0.34%	0.055%
σ	3.3%	0.039%

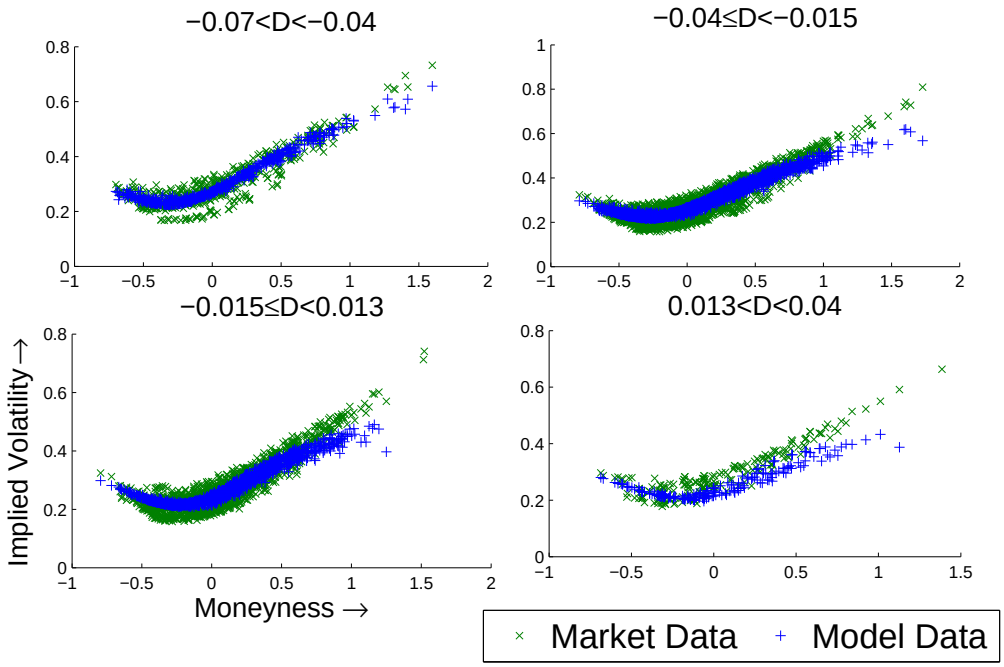

Figure 3.5: Market and model implied volatilities as a function of moneyness. Here we have four plots where we have separated data based on the value of the offset function at time of quotation. Vol_0 is used in the above model data with $\lambda = 4$.

function of a single variable. This is useful as our dataset comprises of options over a range of maturities and underlying asset values. Our modelling assumption is that option prices are functions of the spot price P_t, interest rate $r(t)$, time to maturity $(T-t)$, strike K and offset value D_t. Since the moneyness variable uniquely incorporates the set of variables $(P_t, r(t), T-t, K)$, the remaining variation, or spread in implied volatility can only be due to the effect of D_t. A quick visual check shows

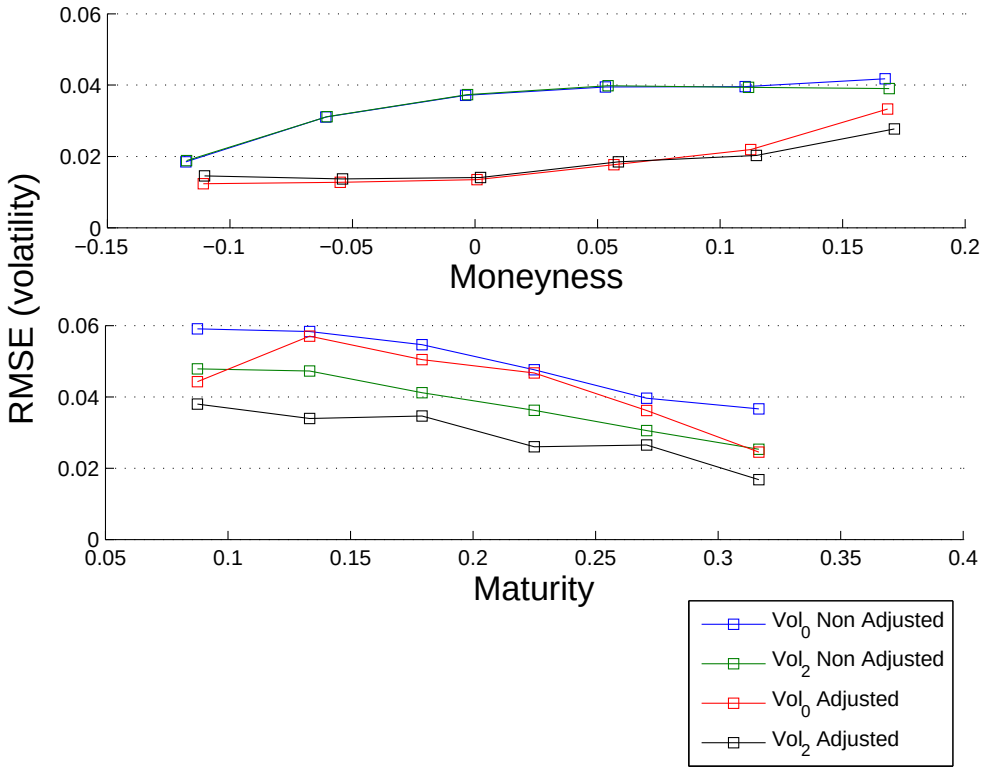

Figure 3.6: Plots of the RMSE (volatility) as a function of time to maturity and moneyness.

that the implied volatility spread in market data is larger than that of the model data, especially when the value of the offset function is close to zero. There are two possibilities here to explain this difference. Either our modelling assumption i.e. the spread in implied volatility is explained by the value of D_t is inaccurate, or, there is another explanatory variable that we are not taking into account. Later on we consider the possibility of a 'calendar' spread i.e. a spread due to the calendar time dependence of the market quotes/market implied volatilities. Taking account of this 'calendar' spread dramatically reduces our error and we describe this approach in Section 3.7.2.

We can also look at the error as a function of time to maturity and moneyness as is

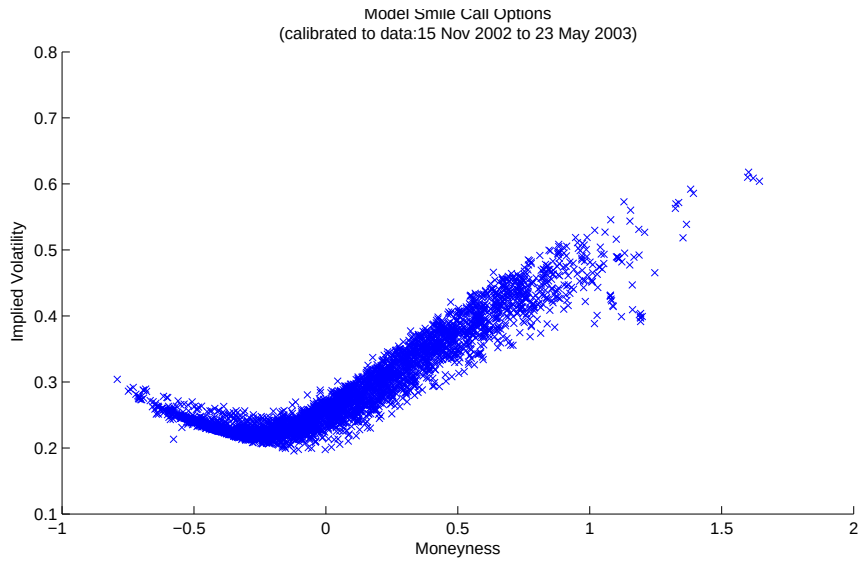

Figure 3.7: Here we plot model Black implied volatilites as a function of moneyness.

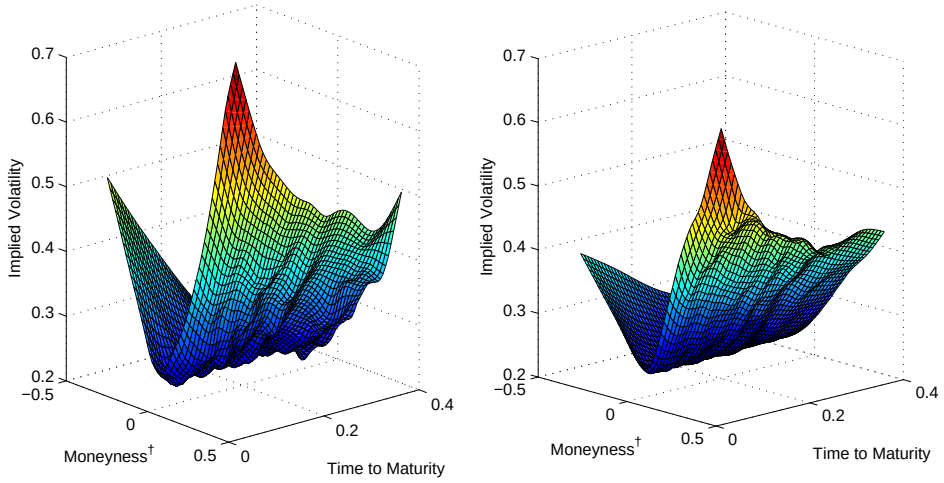

Figure 3.8: Here we plot market implied volatilities and model implied volatilities. For each option, we take the underlying asset value, strike price, time to maturity and value of offset and calculate the Black implied volatility based on the model price given those parameters. † Since we have a time axis here we define moneyness in this case to be $\log(e^{r(T-t)}S/K)$

done in [22]. See Figure 3.6. In this figure we are looking at four datasets. One for each of Vol_0 and Vol_2, and within those, errors for the original dataset and adjusted dataset. Again, the reasons for adjusting our market data are discussed in Section 3.7.3. In general it can be seen that Vol_2 is marginally preferable if we bucket the errors per moneyness, and consistently better than Vol_0 if we bucked the errors per

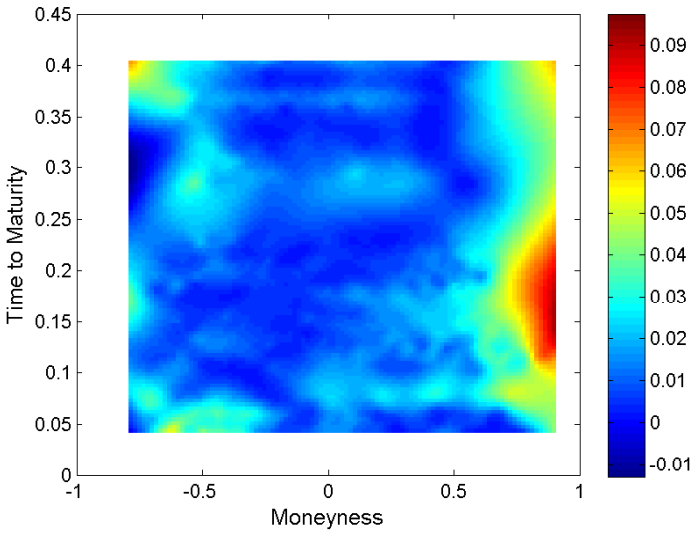

Figure 3.9: Here we make a color map of the absolute value of differences between market and model implied volatilities. We see the largest errors for deep into the money and out of the money options. In general, deep in the money options are overvalued by the model. Blue represents agreement between model and market, while red indicates disagreement. The scale is in implied volatility.

maturity. There are a few points to note here. We see that there is a large variation across the range of moneyness and maturity for each dataset. For short term options, in the case of calibration to non-adjusted data, the error is of the order of 6% while for long term options (0.5 years) we see an error of 4%. Note that the error units here are actual implied volatilities. We see also that the RMSE of volatility grows with moneyness.

Another approach to see if our model is working is to see if the model reproduces the smile seen in market data. In [24] it is reported that the Hobson and Rogers model fails to produce the 'smile' effect. In contrast, our results show good replication of the market smile. See Figure 3.7. Figure 3.8 shows the market smiles and the model smiles for our dataset. In this graph we have interpolated the market and model data and drawn a best fit surface to fit this data. Figure 3.9 shows a top down view of a best fit surface generated from the absolute value of differences between market and implied volatilities as a function of moneyness. The model is most accurate for at the money options, with deep in the money options tending to be overpriced.

In order to improve accuracy of the model we suggest the use of $D_t^{(2)}$, the second

order offset function. The second order offset function has not previously been used in the literature and no attempt has been made to validate this type of model. We tackle this problem in Chapter 4.

A major problem with the calibration method is the choice of a good set of starting parameters. The non-linear least squares function requires an initial set of parameters $(\alpha_1, \alpha_2, \dots)$ with which to being its search of the optimal parameters. It was found that a bad initial choice would lead to convergence to a local minimum of the RMSE which could be an order of magnitude greater than the global minimum. This problem could only be overcome through extensive manual testing of different starting vectors. It may be the case that in the calibration of Vol_1 and Vol_3 we were unable to find such a set of good starting parameters which then resulted in the failure of the calibration routine.

3.7.2 Uniqueness of Implied Volatility

As mentioned earlier, for a given moneyness, there is a large spread in implied volatility values in the market data. This might suggest an opportunity for arbitrage. See Figure 3.10. Closer examination of the data shows this not to be the case. We find that the implied volatility surface is a function of time of quotation. The spread

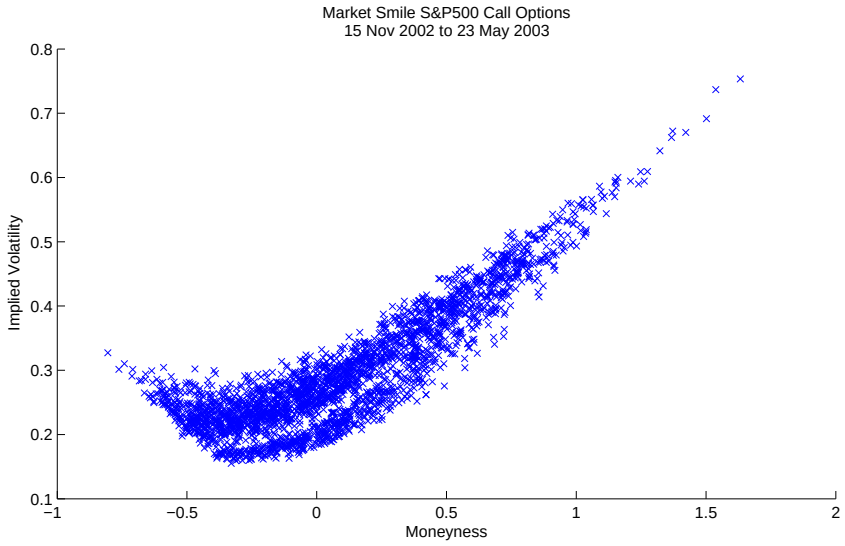

Figure 3.10: Implied Volatility as a function of moneyness

of implied volatilities in the dataset is explained by making the following observation about the dataset: the data consists of prices of options with fixed maturity dates, recorded over some time period. For instance, a similar data set may be constructed by recording the prices, on a particular day, of options due to expire in 30 days, 60 days and 90 days, then on each consecutive day for 30 days, recording the prices of these specific options. After this period of only 30 days the dataset will consist of option prices with expiry dates ranging from 1 to 90 days. In summary, our dataset which we use to calibrate our model, actually consist of different options whose prices are recorded from 15th November 2002 to 23rd May 2003.

In an effort to understand the variation in implied volatilities for a given moneyness and maturity we looked at option prices that were recorded from 15th November to 3rd December 2002. We can see from Figure 3.11, four bands of prices. We can see

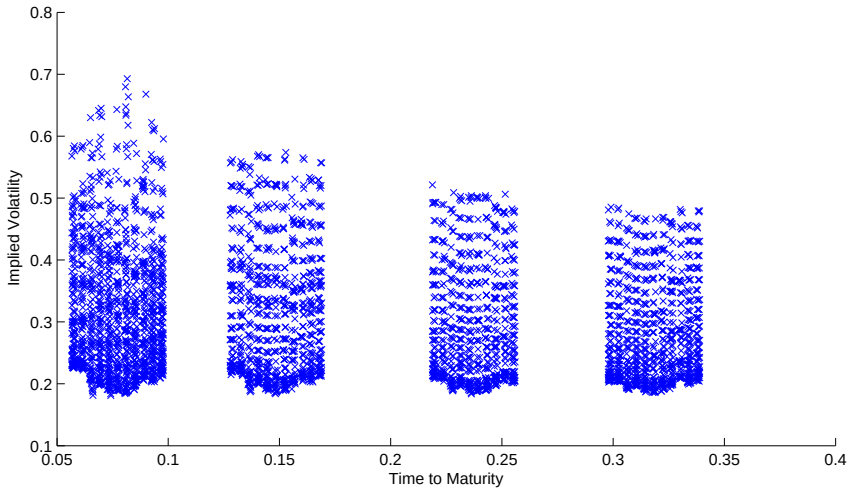

Figure 3.11: Implied Volatility as a function of maturity only. Note that the banded structure is a result of the dataset being constructed from different (in terms of calendar date of maturity) options. See Figure 3.13 to view this graph after the correction for time dependence has been made.(Marked implied volatilities recorded from 15th November to 3rd December 2002.)

that the dataset consists of 4 options recorded over this time period. The right of each band consists of the oldest recorded prices, moving to newer prices to the left. Within each band we can see that the range of implied volatilities for a given maturity vary with the time/date of quotation. The variation due to time affects each

band/option equally. We can make the obvious conclusion from this visual analysis that the reason we are seeing non-unique implied volatilities in the market prices is that the market implied volatilities are a function of time. This may be due to changing market sentiment or some other external factor. We do not have a time parameter in our volatility specification and this will surely damage the validity of the calibration procedure.

Effectively, up until now, we have been trying to calibrate our model to a surface that changes over time which will clearly impair our results. Unfortunately our model can't account for this as there is no explicit time dependence parameter. A better approach is to somehow take account of this time dependence by adjusting our data before calibration. Now when we use the model to determine option prices we will need to readjust back using a 'daily adjustment index'. A method is given in the next section.

3.7.3 Daily adjustment index

A simple approach to resolving the above issue is to 'correct' market prices, according to some daily volatility index. This index should reflect overall shifts in market option prices from one day to the next. We define this index to be the average market implied volatility on a particular day, for each day in the data set. The adjusted option price dataset is then used to calibrate the model. The model would then produce option prices, which would then have to be converted back to the real world prices, again by use of the volatility index.

In order to adjust the market data we have to determine for a given change in volatility, the corresponding necessary change in option price i.e. $\frac{\partial V}{\partial \sigma}$. To determine the change in V for a given change in σ we expand the Black Scholes closed form option price $V(\sigma)$ using a Taylor series, holding all the other parameters constant. We

have

$$\begin{aligned} V(\sigma+\Delta\sigma) &= V(\sigma)+V'(\sigma)\Delta\sigma+\frac{1}{2!}V''(\sigma)(\Delta\sigma)^2 \\ &\quad +\frac{1}{3!}V'''(\sigma)(\Delta\sigma)^3+\mathcal{O}((\Delta\sigma)^4) \end{aligned} \tag{3.16}$$

where

$$\begin{aligned} V'(\sigma) &= P\sqrt{\frac{T-t}{2\pi}}e^{-\frac{d_1^2}{2}}, \\ V''(\sigma) &= P\sqrt{\frac{T-t}{2\pi}}e^{-\frac{d_1^2}{2}}\frac{d_1 d_2}{\sigma} \\ V'''(\sigma) &= P\sqrt{\frac{T-t}{2\pi}}e^{-\frac{d_1^2}{2}}\frac{1}{\sigma^2}\Big[(d_1d_2)^2-d_1^2-d_2^2-d_1d_2\Big] \end{aligned}$$

(Here σ refers to the daily average implied volatility and $\Delta\sigma$ is the difference between the average and the constant volatility as referred to above.)

The adjustment procedure is as follows:

- For each day, determine the Black-Scholes implied volatility for each quoted option price.
- Determine the average implied volatility σ_i for each day i .
- Determine $\Delta\sigma_i$ such that the $\sigma_i + \Delta\sigma_i = const.$ (This constant is arbitrary and we choose $const. = 0.25$ or 25% volatility).
- For each day i, adjust option prices using equation (3.16), substituting σ_i for σ.
  - Now we have a dataset of option prices in which the implied volatility is constant over time.
- Calibrate the model to the adjusted dataset determining the best fit parameters, as described in Section 3.5. The resulting parameters are shown in Tables 3.8 and 3.9.

If we wanted to price a particular option today using the above procedure we would determine the average implied volatility in the market today, and the corresponding

$\Delta\sigma_i$ for the day i on which the option is quoted. Given the parameters for our model $(P_t, D_t^{(1)}, t)$ we refer to the option price surface as output by our model. This surface corresponds to prices of adjusted data and so the price of our option according to this surface is the adjusted price. To determine the real price we effectively invert equation (3.16) to determine $V(\sigma)$ using

$$V(\sigma - \Delta\sigma) = V(\sigma) - V'(\sigma)\Delta\sigma + \frac{1}{2!}V''(\sigma)(\Delta\sigma)^2 - \frac{1}{3!}V'''(\sigma)(\Delta\sigma)^3 + \mathcal{O}((\Delta\sigma)^3)$$

(Here σ refers to the adjusted volatility value and $\Delta\sigma$ refers to the difference between the adjusted volatility value and today's average.) We can see the results of adjusting the volatility surface in two ways. Firstly we compare Figure 3.7.3 and Figure 3.7.3. In each of these figures we have plotted a sample of daily volatility surfaces. It can be seen in the original plot that different days produce very different volatility surfaces, while after the correction this difference is greatly reduced. Secondly in comparing Figure 3.16 and Figure 3.17 we see plots of the volatility smile as calculated using subsets of the original data set corresponding to market data recorded over 10 day intervals. These plots show how the smile changes with the calendar date. We see again that the calendar time variation has been greatly reduced.

Note also that the residual error computed in calibrating the model to the adjusted data can be equated to the residual error when calibrating with the real market data. It can be seen from comparison of the 'unadjusted' results in Tables 3.3 and 3.4, with 'adjusted' results in Tables 3.10 and 3.11, that the residual errors computed using the adjusted data are almost 50% smaller than those of the real data. The results will be discussed in detail in the following section.

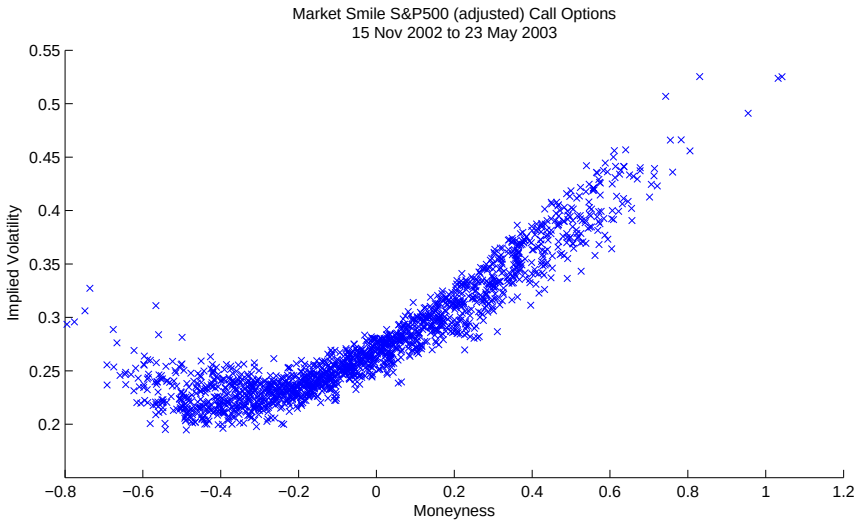

Figure 3.12: Implied Volatility as a function of moneyness. Note that here we use the adjusted prices. The band structure that can be seen in Figure 3.10 has disappeared.

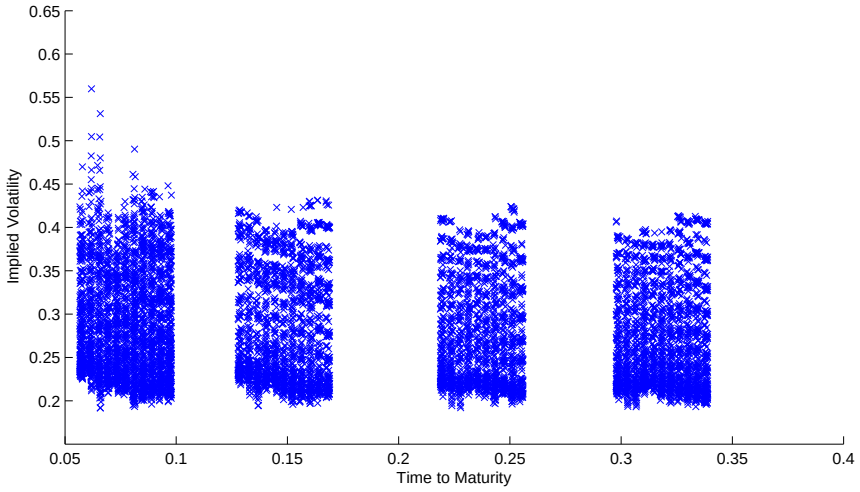

Figure 3.13: This is the corrected data corresponding to Figure 3.11.

3.7.4 Results of calibration with time corrected data.

Tables 3.8 and 3.9 gives the results of the calibration procedure where we have used the adjusted data. Here, the volatility specification will be appended by a superscript 'A'. The results here may be compared with Tables 3.1 and 3.2. Firstly looking at the residual errors associated with these results we see that the residual error for Vol_0 is significantly smaller than that for the non adjusted data, with a similar result for Vol_2. This is to be expected since, by the adjustment process, we have reduced the

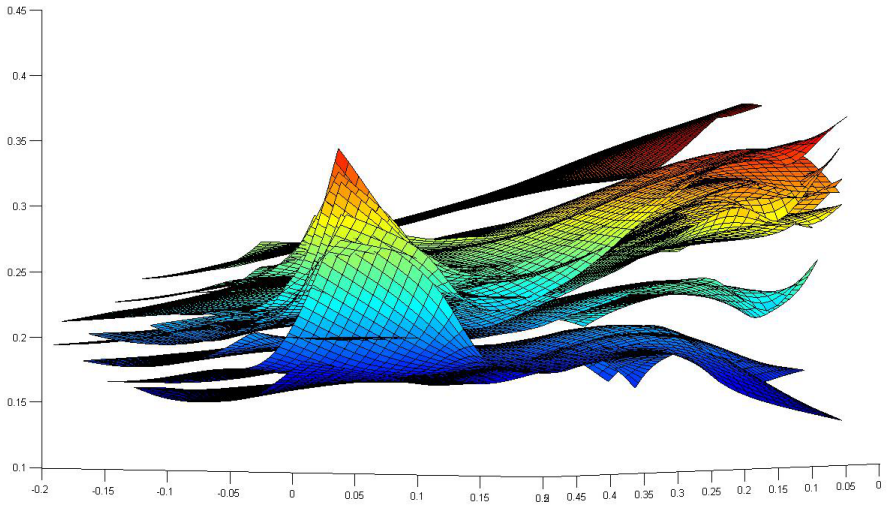

Figure 3.14: Market Implied Volatilities before correction for time dependence.

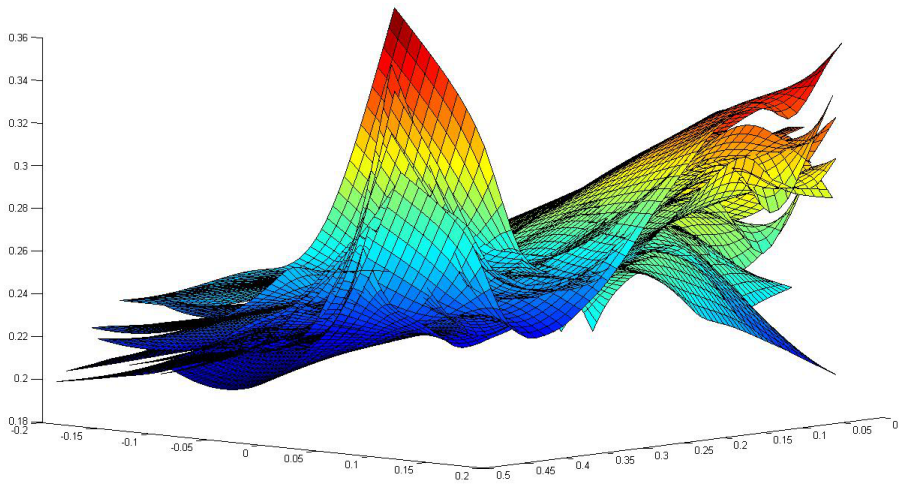

Figure 3.15: Market Implied Volatilities after correction for time dependence: Here we see market implied volatility surfaces. We graph the market implied volatility every 10 trading days from 15th November 2002 to 23rd May 2003.

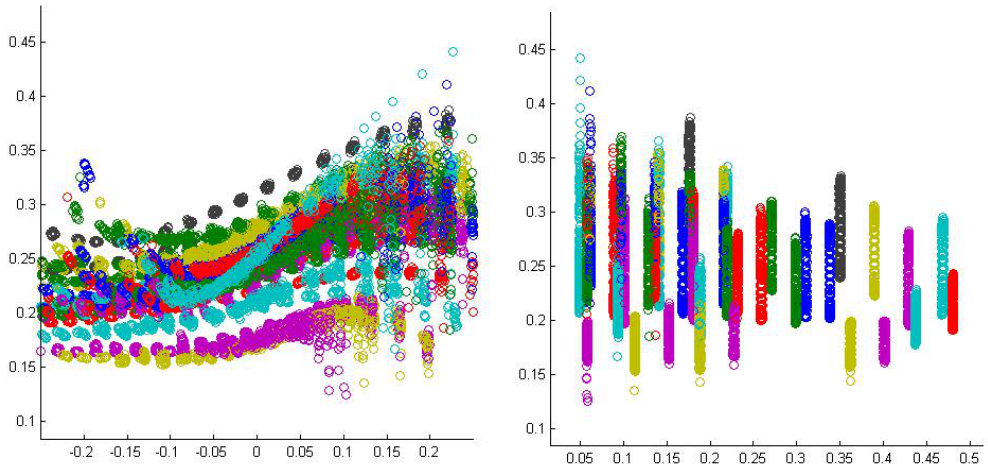

Figure 3.16: We see here the original dataset of market implied volatilities as a function of moneyness (left) and maturity (right). Each colour corresponds to a different day. We see that the average implied volatility does depend on the day the data is taken. For example we can see that the lower band in Figure 3.10 corresponds to the indigo colour in the above graph. The band corresponds to data taken from the month preceding 23rd May i.e. the most recently recorded data only.

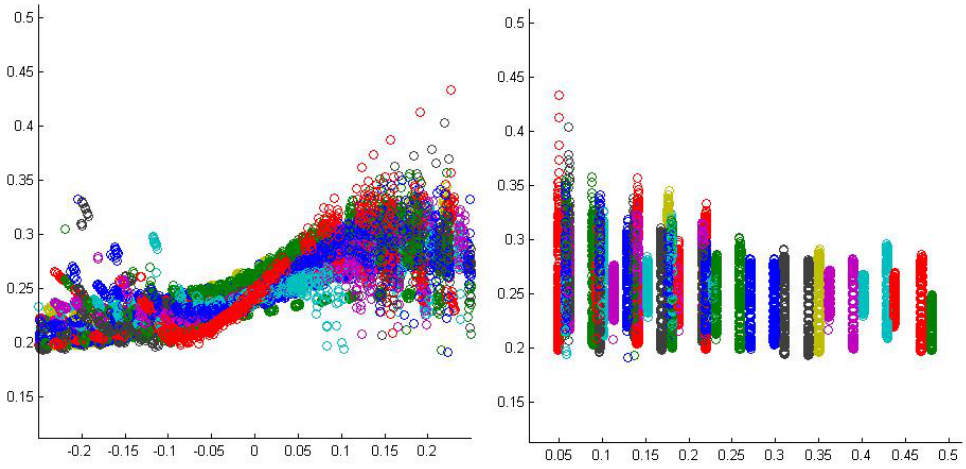

Figure 3.17: Marked implied volatility data before and after daily time dependence 'correction'. Here we see the corrected data. The average market implied volatility is now 0.25. The band structure from Figure 3.16 has disappeared.

spread of implied volatilities for a given moneyness. The adjustment process allows a better calibration to take place. In terms of λ, we are seeing similar values to the previous calibration in the case of Vol_0, and in the case of Vol_2, a slightly higher value of λ. A higher value of λ puts greater weight on more recent market data in calculating the value of the offset. The parameters in the case of Vol_0 are similar for the adjusted and non-adjusted data set. We see much larger difference in the case of Vol_2. The new α_1 parameter is close to zero, and we see α_2 become very dominant. We also see a much larger term in the denominator. If we were to rescale Vol_2 such that the denominator had a value $\simeq 1$ we would find that the α_2 parameter then is of the same order of magnitude to the case where we calibrated to the non-adjusted data. This suggests that the results are in line with the previous calibration. Of course in this case the α_1 parameter disappears completely.

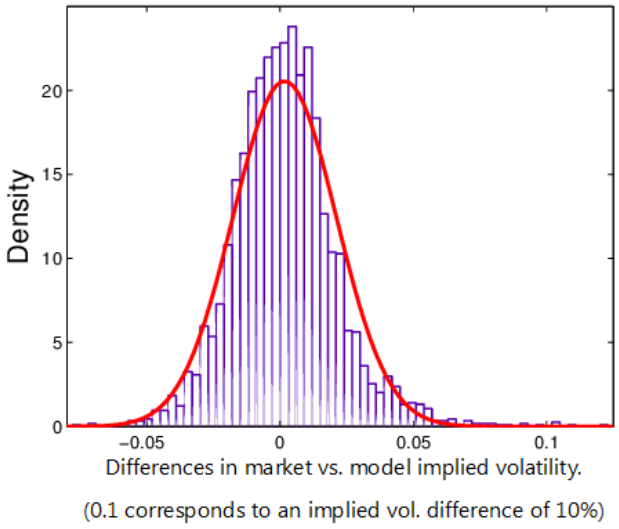

Figure 3.18: Distribution of differences between market implied volatility and model implied volatility in the case of Vol_2, with a fitted normal density function. Here we used the adjusted data set. See Table 3.13 for the mean and standard deviation of this fit.

Table 3.8: Best fit parameters for the $Vol_0^{(A)}$ volatility specification for each of the error specifications.

$Vol_0^{(A)}$	$RMSE^{original}$	$RMSRE(\%)$	$RMSE$	MAE	$MRAE$
λ	4	1	4	4	1
α_1	$0.04 \pm 4.1E-07$	$0.03 \pm 9.6E-07$	$0.04 \pm 4.7E-07$	$0.04 \pm 2.5E-07$	$0.04 \pm 3.0E-08$
α_2	$3.69 \pm 7.4E-05$	$1.26 \pm 1.1E-04$	$3.36 \pm 7.6E-05$	$3.38 \pm 2.9E-05$	$1.78 \pm 4.9E-07$
α_3	$0.09 \pm 9.3E-04$	$0.18 \pm 2.5E-03$	$0.1 \pm 1.0E-03$	$0.1 \pm 6.8E-04$	$0.15 \pm 1.3E-04$

Table 3.9: Best fit parameters for the $Vol_2^{(A)}$ volatility specification for each of the error specifications.

$Vol_2^{(A)}$	$RMSE^{original}$	$RMSRE(\%)$	$RMSE$	MAE	$MRAE$
λ	5	2	5	5	3
α_1	$0.03 \pm 6.6E03$	$0. \pm 4.2E06$	$0. \pm 8.0E04$	$0.01 \pm 5.3E03$	$0. \pm 1.3E05$
α_2	$61.88 \pm 7.0E02$	$80.95 \pm 1.7E05$	$92.17 \pm 5.4E03$	$141.6 \pm 1.2E01$	$159.24 \pm 1.3E00$
α_3	$0.72 \pm 7.1E-01$	$0.8 \pm 1.2E-01$	$0.69 \pm 7.0E-01$	$0.63 \pm 2.0E-01$	$0.67 \pm 7.7E-02$
α_4	$2261.37 \pm 2.2E04$	$2248.34 \pm 4.6E06$	$3005.62 \pm 1.5E05$	$3464.69 \pm 3.1E01$	$3584.02 \pm 2.9E00$
α_5	$-0.43 \pm 2.9E-02$	$-0.55 \pm 1.1E-01$	$-0.43 \pm 4.4E-02$	$-0.46 \pm 9.4E-03$	$-0.51 \pm 5.7E-03$

Table 3.10: Residual error values for each of the error metrics. The volatility specification here is Vol_0^A.

$Vol_0^{(A)}$	$RMSE^{original}$	$RMSRE(\%)$	$RMSE$	MAE	$MRAE$
$RMSE^{original}$	0.0022	0.0029	0.0022	0.0022	0.0023
$RMSRE(\%)$	5.64%	5.38%	5.55%	5.62%	5.37%
$RMSE$	3.28	4.97	3.27	3.29	3.55
MAE	\$1.37	\$1.8	\$1.37	\$1.36	\$1.46
$MRAE$	.0313	.0325	.0309	.0311	.0303

Table 3.11: Residual error values for each of the error metrics. The volatility specification here is Vol_2^A.

$Vol_2^{(A)}$	$RMSE^{original}$	$RMSRE(\%)$	$RMSE$	MAE	$MRAE$
$RMSE^{original}$	0.0021	0.0025	0.0021	0.0021	0.0022
$RMSRE(\%)$	5.45%	5.31%	5.4%	5.49%	5.24%
$RMSE$	2.97	3.95	2.97	2.98	3.13
MAE	\$1.29	\$1.55	\$1.29	\$1.29	\$1.33
$MRAE$	.0299	.0312	.0298	.03	.0295

Table 3.12: Parameters for normal distribution fitted to a distribution of errors between market and model implied volatilities in the case of Vol_0 using the adjusted data set.

Vol_0	Parameter Estimate	Std. Error
μ	0.21%	0.04%
σ	2.5%	0.028%

We can make a similar analysis to that done above with respect to the errors between model and market prices. Once again we see that the model on average overestimates the implied volatility. In the case of Vol_0 we only see a slight improvement with the adjusted data set, but in the case of Vol_2 we see a reduction in the mean error by a factor of 0.5. Naturally, both standard deviation values are much smaller, as we're removed the 'calendar' spread as described above. Visually, we see from Figure 3.18 the model and market prices are matching consistently across all ranges of offset values and the spread between the two is also similar which was the desired effect

Table 3.13: Parameters for normal distribution fitted to a distribution of errors between market and model implied volatilities in the case of Vol_2 using the adjusted data set.

Vol_2	Parameter Estimate	Std. Error
μ	0.16%	0.03%
σ	1.9%	0.022%

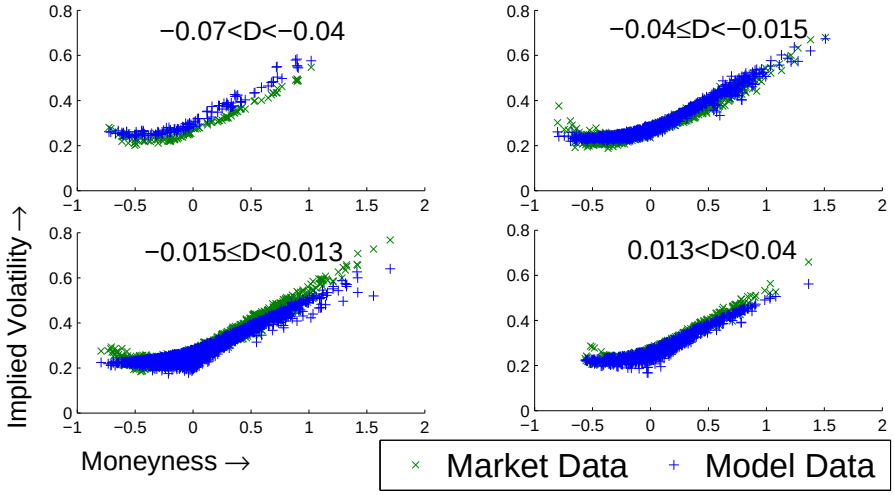

Figure 3.19: Here we plot the market and model implied volatilities for a number of ranges of offset values. This figure may be compared with Figure 3.5. We see better agreement between both datasets across all values of D_t.

from adjusting due to the calendar effect. We can refer back to Figure 3.6 to see the explicit dependence of the RMSE on time to maturity, moneyness and offset. Once again we see that Vol_2 out performs Vol_0 for larger values of moneyness, but in general, the RMSE is consistent across the full range of each of the above parameters which is a welcome result. In terms of maturity, Vol_2 gives significantly better results as was the case for the non-adjusted data. Errors reduce here as maturity increases. We also notice that a smile is still evident in the model implied volatilities.

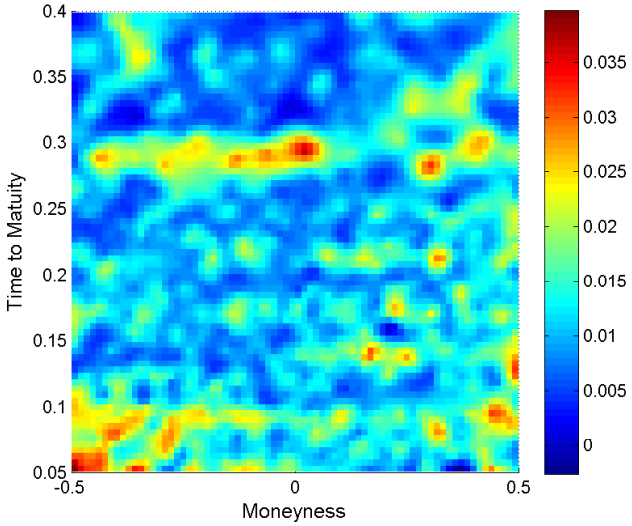

Figure 3.20: Here we make a color map of the differences between market and model implied volatilities where we have calibrated using adjusted data. The scale is in percentage implied volatility.

3.8 Out of sample testing

As a further test, we calibrated our market data sets in the case of Vol_0 andVol_2 with the optimal λ parameters found in the previous calibration, using only market quotations up to 15th November 2002 to 14th February 2003. Using the calibration results we then measured the error when using the calibrated parameters on a different data set, namely data quoted between 15th February 2003 and 23rd May 2003. Table 3.14 gives the results of the initial calibrations, while Figure 3.22 shows the RMSE as a function of moneyness and time to maturity, when the out of sample dataset is used.

We consider first the non adjusted data. Overall we see that the errors have increased as expected. The specific residual data can be found in Table 3.15. In terms of moneyness and maturity we see a similar pattern of error magnitude to that of the calibration to the full market data set, though $Vol_0^{(A)}$ suffers more than the other volatility specifications. The parameters used to generate this data can be found in Table 3.14.

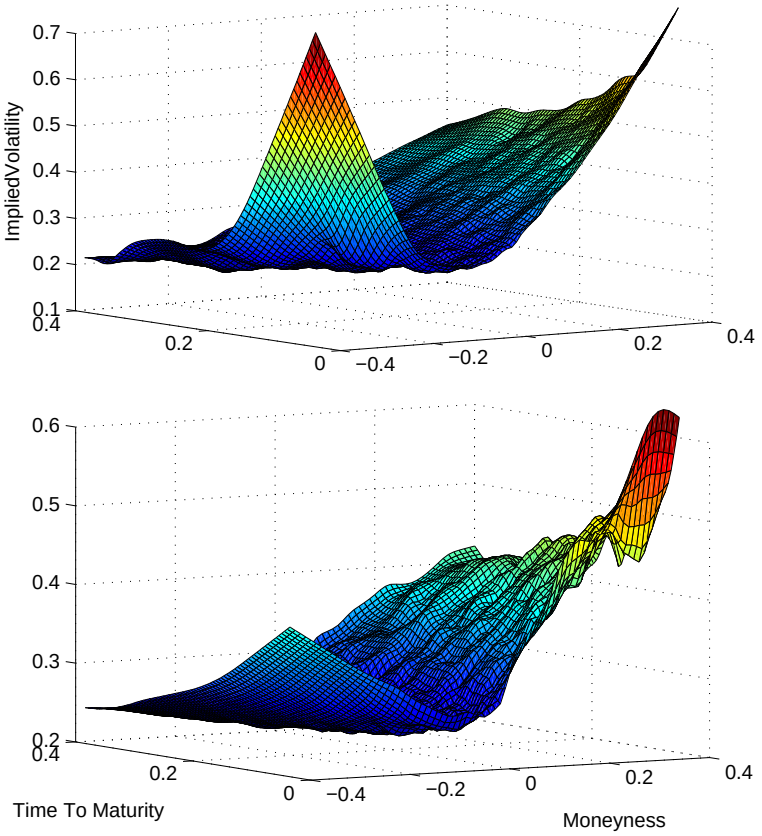

Figure 3.21: Here again we make the visual comparison between the market smile (top) and the model smile (bottom), where we have used the adjusted dataset to calibrate the model. Again we can see that the model reproduces the smile seen in the market data.

3.9 Discussion

In this chapter we have revisited the work of Foschi and Pascucci and implemented a calibration routine using MATLAB code provided by Paolo Foschi. We have made some minor changes to the code as follows: In the implementation we have changed the way in which the value of the offset is computed. We have also changed the method by which the non linear routine finds the optimal fitting parameters.

We examine a number of volatility specifications in addition to the ones suggested in [1] and [22]. We have found that the best choice of volatility functions should be of the form

$$\sigma^2 = \frac{\alpha_1 + \alpha_2(D_t - \alpha_3)^2}{\alpha_4 + \alpha_5(D_t - \alpha_6)^2},$$

Table 3.14: Table of calibrated parameters for a subset of original market data set. Data corresponds to 15th November 2002 to 14th February 2003. These parameters are then used in the out-of-sample testing.

	Vol_0	Vol_2	$Vol_0^{(A)}$	$Vol_2^{(A)}$
λ	1	1	1	1
α_1	0.04	250	0.037	15.69
α_2	2.1	6216.1	1.68	11.64
α_3	0.1326	0.132	0.145	0.33
α_4		5632.82		1590.6
α_5		-0.920		-0.3861

Table 3.15: Results from out of sample testing. In these tests the error metric used was RMSRE. This table presents the results from each of the volatility specifications.

	Vol_0	Vol_2	$Vol_0^{(A)}$	$Vol_2^{(A)}$
MAE	\$3.88	\$4.09	\$1.804	\$1.74
$RMSE^{original}$	0.00595	0.00631	0.00281	0.00286
$RMSRE(\%)$	19.23%	18.99%	8.22%	7.30%
$RMSE$	27.07	29.45	5.41	5.34
$MRAE$	0.104	0.105	0.0433	0.0408

a form which was suggested first in [25].

For the first time we examine the dependence of the model on λ and conclude that the optimal choice of λ depends on the volatility specification being used.

We calibrate using the dataset used in [22] but note the calendar time dependence of the volatility smile, independent of other underlying variables. We make adjustments to the dataset in order to minimise this effect and implement the calibration routine, again testing all volatility specifications on this new dataset. The model fit is much better on this adjusted dataset and this improvement can be translated back into the calibration of real market data.

In the non adjusted dataset we find the RMSE of implied volatility for Vol_0 and Vol_2 to be $\simeq 12.6\%$ while in the adjusted data set we find an error of $\simeq 5.6\%$ for Vol_0 and $\simeq 5.4\%$ for Vol_2. The results suggest that significant improvements can still be made to the model. These improvements will be described in the next chapter. Firstly we will look at using the second offset function $D_t^{(2)}$ and consider our option and

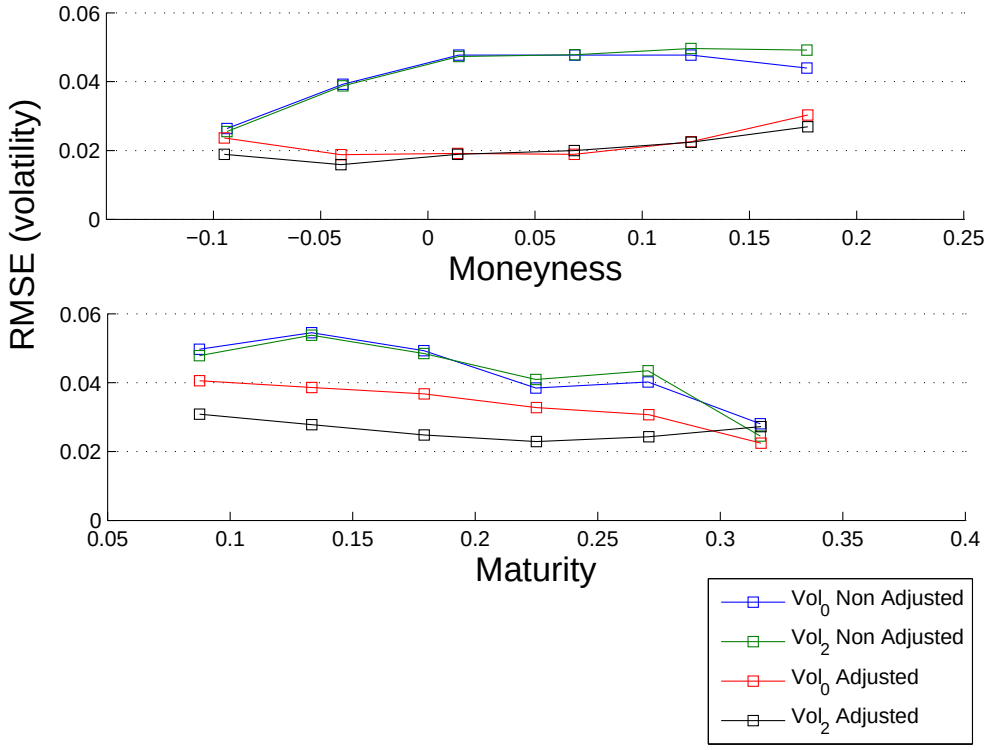

Figure 3.22: Plots of RMSE as a function of time to maturity, moneyness and offset values, for the out-of-sample market data set. Here we compare errors for Vol_0 and Vol_2 using adjusted and non adjusted market data sets. All market date used here dates from 15th February 2003 to 23 March 2003. The model data here is generated using parameters calibrated to data quoted between November 2002 and 15th February 2003.

volatility to be a function of both $D_t^{(1)}$ and $D_t^{(2)}$. The second offset function has the advantage that it doesn't allow for cancellation. (due to the squared term). In considering $D_t^{(2)}$ we have to determine a new partial differential equation and implement a new higher order numerical procedure. Secondly we will investigate more closely the role of λ and consider a variable λ and a generalised offset function.

Chapter 4

Calibration of a more general model

4.1 Introduction

In this chapter we aim to extend the Hobson and Rogers model to include dependence on both first and second order offset functions. It has been shown in Chapter 3 and noted in other articles [24] that the smiles produced by the Hobson Rogers model are 'too shallow'. We also note from empirical tests carried out by us that a quadratic dependence of the volatility on the first offset function alone may not be suitable. Finally, we justify the introduction of the second order offset function into our model by noting that the first order offset function loses price information in the summation due to cancellation. This is overcome with using $m = 2$. Here we derive a partial differential equation which describes the price of a European Call option under the new assumption. We then describe how to extend the numerical scheme as described in Chapter 3 to numerically solve this equation.

4.2 Derivation of the p.d.e.

Proposition 20. *Under the Hobson and Rogers model, the price of a European call option with strike price K, maturity T, underlying risk free rate r, and dependence on the first and second order offest functions, namely $D_t^{(1)}$ and $D_t^{(2)}$, obeys the following partial differential equation*

$$\begin{aligned} & rPf_P - \lambda D_t^{(1)} f_{D^{(1)}} - \lambda D_t^{(2)} f_{D^{(2)}} - f_t \\ +\sigma^2\Bigg(& -(D_t^{(1)}-1) f_{D^{(2)}} + P f_{PD^{(1)}} - \frac{1}{2}\Big(f_{D^{(1)}} + P f_{PP} + f_{D^{(1)}D^{(1)}}\Big) \\ & +2\Big((D_t^{(1)})^2 f_{D^{(2)}D^{(2)}} + P f_{PD^{(2)}} + D_t^{(1)} f_{D^{(1)}D^{(2)}}\Big)\Bigg) \;=\; 0 \end{aligned} \tag{4.1}$$

with boundary conditions

$$f(P_T, D_T^{(1)}, D_T^{(2)}, 0) = \max(P_T - K, 0).$$

Proof. We use the result in [1] to write down a stochastic differential equation for $D_t^{(m)}$, namely

$$dD_t^{(m)} = mD_t^{(m-1)} dW_t^{\mathbb{P}} + \frac{m(m-1)}{2} D_t^{(m-2)} \langle dZ_t \rangle^2 - \lambda D_t^{(m)} dt,$$

where $W_t^{\mathbb{P}}$ is some $\mathbb{P}$-Wiener process. Now for $m = 2$ we have

$$\begin{aligned} dD_t^{(2)} &= 2D_t^{(1)} dZ_t + (\sigma^2 - \lambda D_t^{(2)}) dt \\ &= 2D_t^{(1)}\Big(\mu(D_t^{(1)}, D_t^{(2)}) dt + \sigma(D_t^{(1)}, D_t^{(2)}) dW_t^{\mathbb{P}}\Big) + (\sigma^2 - \lambda D_t^{(2)}) dt \\ &= \Big(2\mu D_t^{(1)} + \sigma^2 - \lambda D_t^{(2)}\Big) dt + 2\sigma D_t^{(1)} dW_t^{\mathbb{P}}. \end{aligned} \tag{4.2}$$

We wish to write this expression using the risk neutral measure $\mathbb{Q}$ as derived in the previous chapter, namely

$$dW_t^{\mathbb{P}} = \varphi_t dt + dW_t^{\mathbb{Q}},$$

where

$$\varphi(t) = -\frac{1}{2}\sigma(D_t^{(1)}, D_t^{(2)}) - \mu(D_t^{(1)}, D_t^{(2)})/\sigma(D_t^{(1)}, D_t^{(2)}).$$

Using this measure we rewrite equation (4.2) as

$$\begin{aligned}
dD_t^{(2)} &= \Big(2\mu D_t^{(1)} + \sigma^2 - \lambda D_t^{(2)}\Big)dt + 2\sigma D_t^{(1)} dW_t^{\mathbb{P}} \\
&= \Big(2\mu D_t^{(1)} + \sigma^2 - \lambda D_t^{(2)}\Big)dt + 2\sigma D_t^{(1)}\Big(\varphi_t dt + dW_t^{\mathbb{Q}}\Big) \\
&= \Big(2\mu D_t^{(1)} + \sigma^2 - \lambda D_t^{(2)}\Big)dt + 2\sigma D_t^{(1)}\left(\Big(-\frac{1}{2}\sigma - \frac{\mu}{\sigma}\Big)dt + dW_t^{\mathbb{Q}}\right) \\
&= \Big(2\mu D_t^{(1)} + \sigma^2 - \lambda D_t^{(2)} - \sigma^2 D_t^{(1)} - 2D_t^{(1)}\mu\Big)dt + 2\sigma D_t^{(1)} dW_t^{\mathbb{Q}} \\
&= -\Big(\lambda D_t^{(2)} + \sigma^2 D_t^{(1)} - \sigma^2\Big)dt + 2\sigma D_t^{(1)} dW_t^{\mathbb{Q}}.
\end{aligned}$$

To remind the reader, we write expression for dP and $dD_t^{(1)}$ under the risk neutral measure as follows:

$$\begin{aligned}
dP_t &= rP_t dt + \sigma P_t dW^{\mathbb{Q}}, \\
dD_t^{(1)} &= -\left(\frac{1}{2}\sigma^2 + \lambda D_t^{(1)}\right)dt + \sigma dW^{\mathbb{Q}}.
\end{aligned}$$

We now use the above results and the Feynman-Kac formula to write a p.d.e. that describes the option price $f(P_t, D_t^{(1)}, D_t^{(2)}, T-t)$. We have

$$\frac{\partial f}{\partial t} + \mathcal{A}f = 0, \tag{4.3}$$

where

$$\mathcal{A}f(t,x) = \sum_{i=1}^{4} \mu_i \frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j=1}^{4} C_{i,j}(t,x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x),$$

where

$$C_{i,j}(t,x) = \sigma_i \sigma_j.$$

We let the indices $i = 1, 2, 3$ correspond to $(P_t, D_t^{(1)}, D_t^{(2)})$ respectively. Equation (4.3)

written out term by term becomes,

$$\begin{aligned}
rP\frac{\partial f}{\partial P} - \left(\frac{1}{2}\sigma^2 + \lambda D_t^{(1)}\right)\frac{\partial f}{\partial D^{(1)}} - \left(\lambda D_t^{(2)} + \sigma^2 D_t^{(1)} - \sigma^2\right)\frac{\partial f}{\partial D^{(2)}} \\
+\frac{1}{2}\sigma^2 P^2\frac{\partial^2 f}{\partial P^2} + \frac{1}{2}\sigma^2 P\frac{\partial^2 f}{\partial P \partial D^{(1)}} + \frac{1}{2}2\sigma^2 P D_t^{(1)}\frac{\partial^2 f}{\partial P \partial D^{(2)}} \\
+\frac{1}{2}\sigma^2 P\frac{\partial^2 f}{\partial D^{(1)}\partial P} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial (D^{(1)})^2} + \frac{1}{2}2\sigma^2 D_t^{(1)}\frac{\partial^2 f}{\partial D^{(1)}\partial D^{(2)}} \\
+\frac{1}{2}2\sigma^2 P D_t^{(1)}\frac{\partial^2 f}{\partial D^{(2)}\partial P} + \frac{1}{2}2\sigma^2 D_t^{(1)}\frac{\partial^2 f}{\partial D^{(2)}\partial D^{(1)}} + \frac{1}{2}(2\sigma D_t^{(1)})^2\frac{\partial^2 f}{\partial (D^{(2)})^2} - \frac{\partial f}{\partial t} &= 0.
\end{aligned}$$

Upon factorising and simplifying the notation we have

$$\begin{aligned}
rPf_P - \lambda D_t^{(1)} f_{D^{(1)}} - \lambda D_t^{(2)} f_{D^{(2)}} - f_t \\
+\sigma^2\Bigg(-\frac{1}{2}\Big(f_{D^{(1)}} + Pf_{PP} + f_{D^{(1)}D^{(1)}}\Big) \\
-(D_t^{(1)} - 1)f_{D^{(2)}} + Pf_{PD^{(1)}} \\
+2\Big((D_t^{(1)})^2 f_{D^{(2)}D^{(2)}} + Pf_{PD^{(2)}} + D_t^{(1)} f_{D^{(1)}D^{(2)}}\Big)\Bigg) &= 0. \qquad (4.4)
\end{aligned}$$

The boundary condition comes from the definition of the European Call option and is given by $f(P_T, D_T^{(1)}, D_T^{(2)}, 0) = \max(P_T - K, 0)$. □

To implement a numerical scheme for the above equation would seem difficult so, as done in the paper by Foschi and Pascucci [22], we try to find a transformation of the variables $(f, P, D_t^{(1)}, D_t^{(2)}, t)$ to reduce the above p.d.e. to a simpler form.

Proposition 21. *By use of the transformation*

$$\begin{aligned}
f(P, D_t^{(1)}, D_t^{(2)}, T-t) &\rightarrow Ke^{r(T-t)}V(x, y, z, \tau), \\
x &= \log(\frac{P_t}{K}) - r(T-t), \\
y &= \log(\frac{P_t}{K}) - r(T-t) - D_t^{(1)}, \\
z &= (D_t^{(1)})^2 - \frac{D_t^{(2)}}{2} \quad \textit{and} \\
\tau &= -\lambda(T-t),
\end{aligned}$$

the partial differential equation describing the price of a European call option as given

in equation (4.1) may be written as

$$\frac{\sigma^2}{2\lambda}(V_{xx} - V_x - V_z) + (x-y)V_y - zV_z - V_t = 0. \tag{4.5}$$

with boundary condition

$$V(x_T, y_T, z_T, 0) = (e^{x_T} - 1)^+.$$

Proof. Using the above transformation of variables we have

$$\begin{aligned}
f_P &= e^{r(T-t)}K\left(\frac{V_x}{P} + \frac{V_y}{P}\right), \\
f_{PP} &= e^{r(T-t)}K\left(\frac{1}{P^2}(V_{xx} + V_{xy} - V_x) + \frac{1}{P^2}(V_{yy} + V_{yx} - V_y)\right), \\
f_{D^{(1)}P} &= e^{r(T-t)}K\left(-\frac{1}{P}V_{xy} - \frac{1}{P}V_{yy} + \frac{1}{P}(V_{yz} + V_{xz})\right), \\
f_{D^{(1)}} &= -e^{r(T-t)}K(V_y - yV_z), \\
f_{D^{(1)}D^{(1)}} &= e^{r(T-t)}K(V_{yy} - 2yV_{yz} + y^2V_{yy}), \\
f_{D^{(2)}} &= -\frac{1}{2}e^{r(T-t)}KV_z, \\
f_{D^{(2)}D^{(2)}} &= \frac{1}{4}e^{r(T-t)}KV_{zz}, \\
f_{PD^{(2)}} &= -\frac{1}{2P}e^{r(T-t)}K(V_{yz} + V_{xz}) \quad \text{and} \\
f_t &= e^{r(T-t)}K\left(-rV + rV_x + rV_y + \lambda V_t\right),
\end{aligned}$$

which, when substituted into equation (4.4) gives

$$\frac{\sigma^2}{2\lambda}(V_{xx} - V_x - V_z) + (x-y)V_y - zV_z - V_t = 0. \tag{4.6}$$

As usual, the boundary condition $f(P_T, D_T^{(1)}, D_T^{(2)}, 0) = \max(P_T - K, 0)^+$ may be written under the above transformation as

$$V(x_T, y_T, z_T, 0) = (e^{x_T} - 1)^+ \qquad \square$$

Note here that the functional form of $\sigma = \sigma(D_t^{(1)}, D_t^{(2)})$ has not been specified. We

also note that the form of this equation is very similar to the previous case where we only consider dependence on $D_t^{(1)}$,i.e.

$$\frac{\sigma(D_t)^2}{2\lambda}(V_{xx} - V_x) + (x - y)V_y - V_t = 0.$$

As a result, the implementation of the numerical scheme for solving this new p.d.e. is similar to that of the previous one.

4.3 Numerical Scheme

The numerical scheme is broadly similar to the one described in Chapter 3, except now that we have an extra dimension. The parts $\frac{\sigma^2(D_t)}{2\lambda}(V_{xx} - V_x - V_z) - zV_z$ of and $Y(V) := (x - y)V_y - V_t$ of (4.5) are treated separately. We consider the uniform grid

$$G = \Big\{(i\Delta_x, j\Delta_y, k\Delta_z, n\Delta_t) \quad | \quad i, j, k, n \in \mathbb{Z}, n \geq 0\Big\}.$$

To remind the reader we again define the operators. We use the following central difference approximation for ∂_x

$$\partial_x V(x, y, z, t) \cong D_{\Delta_x} V(x, y, z, t) = \frac{V(x + \Delta_x, y, z, t) - V(x - \Delta_x, y, z, t)}{2\Delta_x}, \tag{4.7}$$

and three-point scheme for ∂_{xx}

$$\begin{aligned} \partial_{xx} V(x, y, z, t) &\cong D^2_{\Delta_x} V(x, y, z, t) \\ &= \frac{V(x + \Delta_x, y, z, t) - 2V(x, y, z, t) + V(x - \Delta_x, y, z, t)}{\Delta_x^2}. \end{aligned} \tag{4.8}$$

We combined the V_y and V_t operators as in Chapter 3 so that

$$(x-y)V_y - V_t \cong Y_{\vec{u}}^- V(x,y,z,t) \tag{4.9}$$

$$= \lim_{\Delta_t \to 0} \frac{\tilde{V}((x, y+\Delta_t(x-y), z, t-\Delta_t) - V(x,y,z,t)}{\sqrt{1+(x-y)^2}\Delta_t}, \tag{4.10}$$

In the above approximations

$$\tilde{V}(x,y,z,t) = (1-\gamma)V(x,\tilde{y},z,t) + \gamma V(x,\tilde{y}+\Delta_y,z,t),$$

i.e. we linearly interpolate the values $V(x,y,z,t)$ that are not on the grid. We have $\tilde{y} = \lfloor y/\Delta_y \rfloor \Delta_y$, with $\lfloor \cdot \rfloor$ denoting the integer part, and $\gamma = (y-\tilde{y})/\Delta_y$. We also have the new operator

$$\partial_z V(x,y,z,t) \cong D_{\Delta_z} V(x,y,z,t) = \frac{V(x,y,z+\Delta_z,t) - V(x,y,z-\Delta_z,t)}{2\Delta_z}. \tag{4.11}$$

Moving to more compact notation now, we denote as before

$$V^n_{i,j,k} = V(i\Delta_x, j\Delta_y, k\Delta_z, n\Delta_t), \quad i,j,k \in \mathbb{Z}, \quad |i| \leq I, \quad |j| \leq J, k \leq K, n \leq N.$$

Here I, J, K, N specify the grid size for a given grid spacing. Numerically, we chose I, J, K, N and calculated the corresponding $\Delta_x, \Delta_y, \Delta_z, \Delta_t$. The boundaries of the grid correspond to $\pm I\Delta_x, \pm J\Delta_y, [0, K\Delta_z], [0, N\Delta_t]$. (In the numerical scheme, $t = 0$ corresponds to expiry.). Applying the above notation to the operator $D^2_{\Delta_x} - D_{\Delta_x}$ for $|i| \leq (I-1)$ gives

$$D_{\Delta_x^2} V^n_{i,j,k} - D_{\Delta_x} V^n_{i,j,k} = (d_1 V^n_{i-1,j,k} + d_2 V^n_{i,j,k} + d_3 V^n_{i+1,j,k}),$$

where as before, $d_1 = 1/\Delta_x^2 - 1/(2\Delta_x)$, $d_2 = -2/\Delta_x^2$ and $d_3 = 1/\Delta_x^2 + 1/(2\Delta_x)$. Similarly, we apply the notation to the operator in (4.9), to obtain

$$\begin{aligned}\frac{\tilde{V}\Big((x, y + \Delta_t(x-y), z, t-\Delta_t\Big) - V\Big(x,y,z,t\Big)}{\sqrt{1+(x-y)^2\Delta_t}} &= d_4\Big\{(1-\gamma)V\Big(x, y+s\Delta_y, z, t-\Delta_t\Big) \\ &+ \gamma V\Big(x, y+(s+1)\Delta_y, z, t-\Delta_t\Big) - V\Big(x,y,z,t\Big)\Big\} \\ &= d_4\Big\{(1-\gamma)V^{n-1}_{i,j+s,k} + \gamma V^{n-1}_{i,j+s+1,k} - V^n_{i,j,k}\Big\}\end{aligned}$$

where $d_4 = \Big(1+(x-y)^2\Delta_t\Big)^{-\frac{1}{2}}$ and

$$s = \left\lfloor \frac{(x-y)\Delta_t}{\Delta_y} \right\rfloor = \left\lfloor \Big(i\frac{\Delta_x}{\Delta_y} - j\Big)\Delta_t \right\rfloor \quad \text{and} \quad \gamma = \frac{(x-y)\Delta_t}{\Delta_y} - s.$$

Note that we choose J such that $J \leq \max(j+s+1)$. We return to this in Section 4.4. Finally, for $|k| \leq (K-1)$ we have

$$D_{\Delta_z}V(x,y,z,t) = d_5\Big(V^n_{i,j,k+1} - V^n_{i,j,k-1}\Big),$$

where $d_5 = 1/(2\Delta_z)$. Before we explicitly state the mechanics of the numerical scheme we first provide a rough outline how the scheme will work. We firstly write equation (4.5) in matrix form. We will specify later how to construct the matrices. Let A be the matrix that applies $\sigma^2/2\lambda$ to the relevant terms in our matrix equation. We have

$$\begin{aligned} A.\Big(D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}\Big).V + Y^-_{\Delta_t}.V^n - z.D_{\Delta_z}.V^n &= 0 \\ \Rightarrow A.\Big(D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}\Big).V^n - z.D_{\Delta_z}.V^n &= - Y^-_{\Delta_t}.V^n \\ &= -d_4.M.V^{n-1} + d_4V^n \\ \Rightarrow \Big(\frac{1}{d_4}\Big).A.\Big(D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}\Big).V^n - \Big(\frac{1}{d_4}\Big)z.D_{\Delta_z}.V^n - \mathcal{I}.V^n &= -M.V^{n-1} \\ \Rightarrow \Big\{\mathcal{I} - \Big(\frac{1}{d_4}\Big).A.\Big(D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}\Big) + \Big(\frac{1}{d_4}\Big)z.D_{\Delta_z}\Big\}.V^n &= M.V^{n-1} \\ \Rightarrow \Big\{\mathcal{I} - \Big(\frac{1}{d_4}\Big).A.\Big(D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}\Big) + \Big(\frac{1}{d_4}\Big)z.D_{\Delta_z}\Big\}.V^n &= M.V^{n-1} \end{aligned}$$

$$\Rightarrow V^n = \Big\{\mathcal{I} - \Big(\frac{1}{d_4}\Big).A.\Big(D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}\Big) + \Big(\frac{1}{d_4}\Big)z.D_{\Delta_z}\Big\}^{-1}.M.V^{n-1}$$

where M is a matrix constructed such that $M.V^{n-1}_{i,j,k} = (1-\gamma)V^{n-1}_{i,j+s,k} + \gamma V^{n-1}_{i,j+s+1,k}$. In the above we also have the term $\left(\frac{1}{d_4}\right)$ which is a value calculated at every point in the grid, converted into a matrix form and then applied to the above expressions. $\mathcal{I}$ is the $IJ \times IJ$ identity matrix.

We that the calculation of V^n only requires calculation of a matrix inverse and two matrix multiplications. The difficulty arises in the construction of the above matrices. In the case of the $M.V^{n-1}$ calculation, we have no dependence on the z grid point. For each 'layer' in our three dimensional grid (corresponding to a fixed $k^*\Delta_z$), we store the values of V^{n-1}_{i,j,k^*} in lexicographic form in a column matrix of dimension $I \times J$ by 1, i.e.

$$V^{n-1} = (V^{n-1}_{1,1,k^*}, V^{n-1}_{2,1,k^*}, V^{n-1}_{3,1,k^*}, \ldots, V^{n-1}_{I-1,J,k^*}, V^{n-1}_{I,J,k^*})^T.$$

We then construct M, which doesn't depend on k, such that when we apply $M.V^{n-1}$ (dot product) we get a column matrix of dimension $I \times J$ by 1 whose $i + j(I-1)$th component has value $(1-\gamma)V^{n-1}_{i,j+s,k} + \gamma V^{n-1}_{i,j+s+1,k}$. In the construction of M we must calculate the value s for each i, j. Where the index $i, j+k+1$ refers to a point outside the grid we simply set $j + k + 1 = J$. This setting does not affect our calculations if we choose a large enough grid size.

For example, let us say we choose a grid size equal to J, and spacing Δ_y such that our grid points range from $-j_0\Delta_y$ to $+j_0\Delta_y$. This notation is convenient for the following explanation but as mentioned in Chapter 3, the notation used in the MATLAB script is slightly different due to easier usage. Going back to the current example, we firstly choose a j_0 $(= (J-1)/2)$. Then, based on the range of $D^{(1)}$ values which we would like our grid to cover, we calculate the appropriate Δ_y. Now, as mentioned in Chapter 3, the index in the $y-$direction will vary between $\min(j, j+k)$ and $j+k+1$. On the $+j_0$ edge of our grid the maximum value of k is -0.11. As a result, points on the $+j_0$ boundary will at most depend on the V_{j_0+1} values at the previous time step, and so a dependence on values outside the grid. We will discuss how we deal with this below.

We do not encounter a problem on the $-j_0$ edge as the minimum value of k is $+0.11$ and so only depends on interior grid points. As we move from $-j_0$ to $+j_0$ the contribution from k does not decrease quickly enough to cause a boundary problem and so the numerical scheme does not depend on points outside the grid. As a result, as in Chapter 3, we do not need to impose boundary conditions at the $-j_0$ edge.

On the $+j_0$ edge we will at most have a dependency on points at the index point $j_0 + N$, where N is the number of time steps. Or, after implementing the numerical scheme, we only use grid values in the range $(-j_0, j_0 - N)$ then effectively we are eliminating all grid values that have at some level depended on values of V outside the original grid. If we do this we do not need to impose boundary conditions on the $+j_0$ edge.

Now considering the construction of $\mathcal{I} - \left(\frac{1}{d_4}\right).A.\left(D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}\right) + \left(\frac{1}{d_4}\right)z.D_{\Delta_z}$, we note that we have no dependence the y grid point. For each 'layer' in our three dimensional grid (corresponding to a fixed $j^*\Delta_y$), we store the values of $M.V^{n-1}_{i,j^*,k}$ in lexicographic form in a column matrix of dimension $I \times K$ by 1, i.e.

$$V^{n-1} = (V^{n-1}_{1,j^*,1}, V^{n-1}_{2,j^*,1}, V^{n-1}_{3,j^*,1}, \ldots, V^{n-1}_{I-1,j^*,K}, V^{n-1}_{I,j^*,K})^T.$$

We then construct $\mathcal{I} - \left(\frac{1}{d_4}\right).A.\left(D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}\right) + \left(\frac{1}{d_4}\right)z.D_{\Delta_z}$, which doesn't depend on j, such that when we apply

$$\left\{\mathcal{I} - \left(\frac{1}{d_4}\right).A.\left(D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}\right) + \left(\frac{1}{d_4}\right)z.D_{\Delta_z}\right\}.\left(M.V^{n-1}_{i,j^*,k}\right)$$

(dot product) we get a column matrix of dimension $I \times K$ by 1 whose $i + k(I-1)$th component has value

$$1 - \frac{a_{i,j^*,k}}{\sqrt{1 + \left(i\Delta_x - j\Delta_y\right)^2}}\left[d_1\hat{V}^{n-1}_{i-1,j^*,k} + d_2\hat{V}^{n-1}_{i,j^*,k} + d_3\hat{V}^{n-1}_{i+1,j^*,k} - d_5\left(\hat{V}^{n-1}_{i,j^*,k+1} - \hat{V}^{n-1}_{i,j^*,k-1}\right)\right]$$
$$+ \frac{1}{\sqrt{1 + \left(i\Delta_x - j\Delta_y\right)^2}}k\Delta_z d_5\left(\hat{V}^{n-1}_{i,j^*,k+1} - \hat{V}^{n-1}_{i,j^*,k-1}\right),$$

where $\hat{V}^n = M.V^n$, and $a_{i,j,k}$ is the value of $\sigma(x,y,z)^2/2\lambda$ at the grid point, i.e. the components of the matrix A above. To create $\mathcal{I} - \left(\frac{1}{d_4}\right).A.\left(D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}\right) + \left(\frac{1}{d_4}\right)z.D_{\Delta_z}$ we use the following matrices, all of dimension $IK \times IK$:

$$A = \begin{pmatrix} a_{1,j*,1} & 0 & 0 & 0 & 0 \\ 0 & a_{2,j*,1} & 0 & 0 & 0 \\ 0 & 0 & a_{3,j*,1} & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & a_{I,j*,K} \end{pmatrix}$$

where the entries are stored in a diagonal lexicographic form. The matrix $(1/d_4)$ consists of

$$\frac{1}{\sqrt{1 + \left(1\Delta_x - j^*\Delta_y\right)^2}}$$

on the diagonal, and zeros elsewhere. Each sub matrix in the above pattern is of dimension I^2 and there are K of these. The matrix $D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}$ is of form

$$\begin{pmatrix} \ddots & \ddots & \ddots & 0 & 0 & \ddots & \ddots & \ddots & 0 & 0 & \ddots & \ddots & \ddots & 0 & 0 & 0 & \dots \\ \dots & 0 & d_5 & 0 & 0 & 0 & d_1 & d_2 & d_3 & 0 & 0 & 0 & d_5 & 0 & 0 & 0 & \dots \\ \dots & 0 & 0 & d_5 & 0 & 0 & 0 & d_1 & d_2 & d_3 & 0 & 0 & 0 & d_5 & 0 & 0 & \dots \\ \dots & 0 & 0 & 0 & d_5 & 0 & 0 & 0 & d_1 & d_2 & d_3 & 0 & 0 & 0 & d_5 & 0 & \dots \\ \dots & 0 & 0 & 0 & \ddots & \ddots & \ddots & 0 & 0 & \ddots & \ddots & \ddots & 0 & 0 & \ddots & \ddots & \ddots \end{pmatrix}.$$

The spacings in $D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}$ are calculated using the value of I.

$$z = \begin{pmatrix} 1\Delta_z & 0 & 0 & \dots & & & & & & & 0 \\ 0 & \ddots & 0 & \ddots & & & & & & & \\ 0 & 0 & 1\Delta_z & \ddots & & & & & & & \\ \vdots & \ddots & \ddots & 2\Delta_z & & & & & & & \\ & & & & \ddots & & & & & & \\ & & & & & 2\Delta_z & & & & & \\ & & & & & & 3\Delta_z & & & & \\ & & & & & & & \ddots & & & \\ & & & & & & & & (K-1)\Delta_z & \ddots & \ddots & \vdots \\ & & & & & & & & \ddots & K\Delta_z & 0 & 0 \\ & & & & & & & & \ddots & 0 & \ddots & 0 \\ 0 & & & & & & & & \dots & 0 & 0 & K\Delta_z \end{pmatrix}$$

Each sub matrix in the above pattern is also of dimension I^2 and there are K of these. The matrix corresponding to D_z is easily obtained from the $D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}$ matrix. Once the above matrices have been created and combined as described above, we calculate the inverse of $\mathcal{I} - \left(\frac{1}{d_4}\right).A.\left(D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}\right) + \left(\frac{1}{d_4}\right)z.D_{\Delta_z}$ and apply it to $M.V^{n-1}$ to determine V^n for each j. We finally convert the resulting j, $IK \times 1$ matrices back into a three dimensional grid for later analysis.

4.4 Boundary Conditions

As in Chapter 3, we set the x boundary condition to $D^2_{\Delta_x} - D_{\Delta_x} = 0$. To implement this we modify the matrix $D^2_{\Delta_x} - D_{\Delta_x} - D_{\Delta_z}$ such that rows $\{1, I+1, 2I+1, \dots, K(I-1)+1\}$ and rows $\{I, 2I, \dots, KI\}$ are set equal to zero. We saw in the discussion in the previous section that we do not need to impose boundary conditions on the y boundary if we remember to discard surface values that depend on points external to the grid. In the case of the z boundary, we set $D_{\Delta_z} = 0$. When

looking at the dependence of option prices on $D_t^{(2)}$, our market data doesn't suggest any particular boundary condition. This may be due to lack of market for extreme values of moneyness. Certainly, for options close to being at-the-money, an assumption that $D_t^{(2)}$ is a constant function of moneyness seems to be a fair choice and also the easiest to implement. Also note that this choice recovers the Black-Scholes p.d.e. when we convert back to the standard units of price and time. This boundary condition is implemented by removing the d_5 term from rows 1 to I and rows $(I-1)K+1$ to IK.

4.5 Calibration

We calibrate using a series of volatility specifications using the daily adjusted data. The optimal value of λ is again found by a manual search. We use the following volatility specifications:

$$\begin{aligned}
Vol_1 &:= \alpha_1 + \alpha_2(D_t^{(1)} - \alpha_3)^2 + \alpha_4 D_t^{(2)} \\
Vol_2 &:= \alpha_1 + \alpha_2 D_t^{(2)} \\
Vol_3 &:= \frac{\alpha_1 + \alpha_2(D_t^{(1)} - \alpha_3)^2 + \alpha_4 D_t^{(2)}}{1 + \alpha_5(D_t^{(1)} - \alpha_6)^2 + \alpha_7 D_t^{(2)}} \\
Vol_4 &:= \alpha_1 + \alpha_2(D_t^{(2)} - \alpha_3)^2 \\
Vol_5 &:= \frac{\alpha_1 + \alpha_2 D_t^{(2)}}{1 + \alpha_3 D_t^{(2)}}
\end{aligned}$$

As is done in Chapter 3, we determine the optimal parameter sets α, and optimal λ, that minimises each of the error metrics, RMSE, RMSRE, MRAE, MAE and $\text{RMSE}^{\text{original}}$. We use a grid size of $101 \times 81 \times 21$ with $N = 15$.

4.6 Results

The results obtained are presented in the following tables. We found the residual errors generated by the use of Vol_1 and Vol_3 to be significantly better than that of the other volatility specifications. The errors of these other specifications was of the order of 5 times greater than that of Vol_1 and Vol_3. Focusing then on these volatility specifications, we determined the optimal range of λ to be $3 < \lambda < 5$. Note that both of these volatility specifications are generalisations of Vol_0 and Vol_2 from the previous Chapter. Tables 4.1 and 4.2 below gives the calibrated parameters while Tables 4.3 and 4.4 give the residual errors for each error metric.

In comparison to the residual errors of our models in Chapter 3, we see an improvement. Where in Chapter 3 the minimum RMSRE was 5.31% ($Vol_2^{(A)}$), here we find an error of 3.57% when using Vol_3. Note that while are using the adjusted data set in our calibration, we used untransformed option values when calculating these errors. While in Chapter 3, each of the volatility specifications resulted in an over estimation of implied volatility, typically by $\simeq 0.2\%$, this is not the case here. We see from Tables 4.5 and 4.6 that the errors are centered around zero with the average error for Vol_1 being -0.019% and that of Vol_3 0.025%. While this is an order of magnitude reduction, a better measure of model validity is the standard deviation of the errors. We see here that the standard deviation is 2.2% and 1.9% for Vol_1 and Vol_3 respectively. This compares with 2.5% and 1.9% for the corresponding volatility specifications in Chapter 3. Analysing the RMSE of implied volatility as a function of moneyness and time to maturity, we see from Figure 4.1 that Vol_3 performs no better than Vol_1 in terms of moneyness, but offers better performance consistently when we compare options of equal time to expiry.

It is notable that the calibrated parameters of Vol_1 and Vol_3 are of different orders of magnitude. One explanation of this may be that since $D_t^{(2)}$ is of a higher order than $D_t^{(1)}$, due to the squared term, larger parameters values are required in order to compensate, such that the units of each of these terms match. It would be interesting

to investigate volatility specificaitons in which a $\sqrt{D_t^{(2)}}$ term was used.

We also note that the confidence intervals are quite large. These 95% confidence intervals are calculated using MATLAB's *nlparci()* function which takes the residuals from the calibration routine as an input. In practice we have found that the model is much more sensitive its parameters than these confidence intervals would suggest. We note that in all calibrations the α_7 parameter converges to zero. This is the parameter that scales the $D_t^{(2)}$ term in the denominator of the volatility specificaiton, which suggests this term is unnecessary in our model. This term may contribute to the large confidence intervals that we see. On the other hand we typically see a large coefficient for the corresponding term in the numerator which confirms that the inclusion of the D_t^2 helps to fit the model to the market data.

Table 4.1: Calibrated parameters for Vol_1.

Vol_1	$RMSE^{original}$	$RMSRE(\%)$	$RMSE$	MAE	$MRAE$
λ	3	3	3	4	3
α_1	$0.04 \pm 6.6E-07$	$0.02 \pm 3.2E-06$	$0.04 \pm 5.8E-07$	$0.04 \pm 8.6E-08$	$0.03 \pm 4.1E-07$
α_2	$3.14 \pm 5.3E-05$	$2.03 \pm 1.1E-04$	$3.02 \pm 4.8E-05$	$3.63 \pm 7.3E-06$	$2.46 \pm 1.9E-05$
α_3	$0.11 \pm 9.7E-04$	$0.17 \pm 2.1E-03$	$0.12 \pm 9.4E-04$	$0.1 \pm 3.3E-04$	$0.14 \pm 7.4E-04$
α_4	$0.27 \pm 3.5E-04$	$0.15 \pm 4.0E-04$	$0.19 \pm 4.0E-04$	$0.08 \pm 2.2E-04$	$0.09 \pm 2.0E-04$

Table 4.2: Calibrated parameters for Vol_3.

Vol_3	$RMSE^{original}$	$RMSE(\%)$	$RMSE$	MAE	$MRAE$
λ	4	3	4	4	4
α_1	$0.01 \pm 2.37E02$	$0.01 \pm 4.47E02$	$0. \pm 2.59E03$	$8.56 \pm 3.08E02$	$0. \pm 1.39E05$
α_2	$52 \pm 7.87E02$	$125 \pm 2.58E05$	$64 \pm 8.51E02$	$9779 \pm 5.03E01$	16496 ± 6.99
α_3	$0.42 \pm 1.32E-01$	$0.36 \pm 1.35E-01$	$0.43 \pm 1.83E-01$	$0.46 \pm 2.62E-02$	$0.39 \pm 6.76E-03$
α_4	$22.4 \pm 3.21E02$	$50.36 \pm 1.03E05$	$9.45 \pm 2.46E01$	$952.9 \pm 1.76E01$	$789.97 \pm 3.40E01$
α_5	$285 \pm 3.80E03$	$186 \pm 3.70E05$	$393 \pm 5.07E03$	$71293 \pm 8.33E01$	65950 ± 7.38
α_6	-0.57 ± 0.009	$-0.82 \pm .08$	$-0.56 \pm .102$	-0.54 ± 0.0047	-0.63 ± 0.0069
α_7	$0. \pm 1.88E05$	$0. \pm 2.15E06$	$0. \pm 4.01E05$	$0.33 \pm 5.68E04$	$0.14 \pm 4.50E05$

Table 4.3: Residual errors for Vol_1. Each column provides the results per error metric used in the optimization algorithm, while each row give the value of each error calculated using the optimal parameter set from the associated error metric.

Vol_1	$RMSE^{original}$	$RMSRE(\%)$	$RMSE$	MAE	$MRAE$
$RMSE^{original}$	0.0015	0.0016	0.0015	0.0015	0.0015
$RMSRE(\%)$	3.61%	3.58%	3.56%	3.61%	3.52%
$RMSE$	1.35	1.52	1.34	1.36	1.4
MAE	\$.84	\$.93	\$.84	\$.84	\$.87
$MRAE$	.0195	.0194	.0192	.0195	.0191

Table 4.4: Residual errors for Vol_3.

Vol_3	$RMSE^{original}$	$RMSRE(\%)$	$RMSE$	MAE	$MRAE$
$RMSE^{original}$	0.0014	0.0014	0.0014	0.0014	0.0014
$RMSRE(\%)$	3.47%	3.57%	3.47%	3.47%	3.48%
$RMSE$	1.24	1.28	1.24	1.24	1.25
MAE	\$.79	\$.82	\$.79	\$.79	\$.79
$MRAE$	.0186	.0189	.0186	.0185	.0186

Table 4.5: Implied volatility errors for Vol_1

Vol_1	Parameter Estimate	Std.Error
μ	-0.019%	0.037%
σ	2.2%	0.02%

Table 4.6: Implied volatility errors for Vol_3

Vol_3	Parameter Estimate	Std.Error
μ	0.025%	0.031%
σ	1.9%	0.02%

4.7 Out of sample results

Again we have calibrated each model to a subset of the original market data set and measured the errors against an out of sample data set. While Vol_3 performs better than Vol_1 in the original calibration, that difference has been reversed in the case of the out-of-sample errors. Overall the RMSE values are larger as can be seen from Table 4.8.

Table 4.7: Calibrated parameters for market data observed between November 2002 and 15th February 2003.

	α_1	α_2	α_3	α_4	α_5	α_6	α_7
Vol_1	0.0315	2.64	0.133	$5.739E-10$	n/a	n/a	n/a
Vol_3	3.33	113.168	0.227	29.19	217.66	-0.734	$8.57E-04$

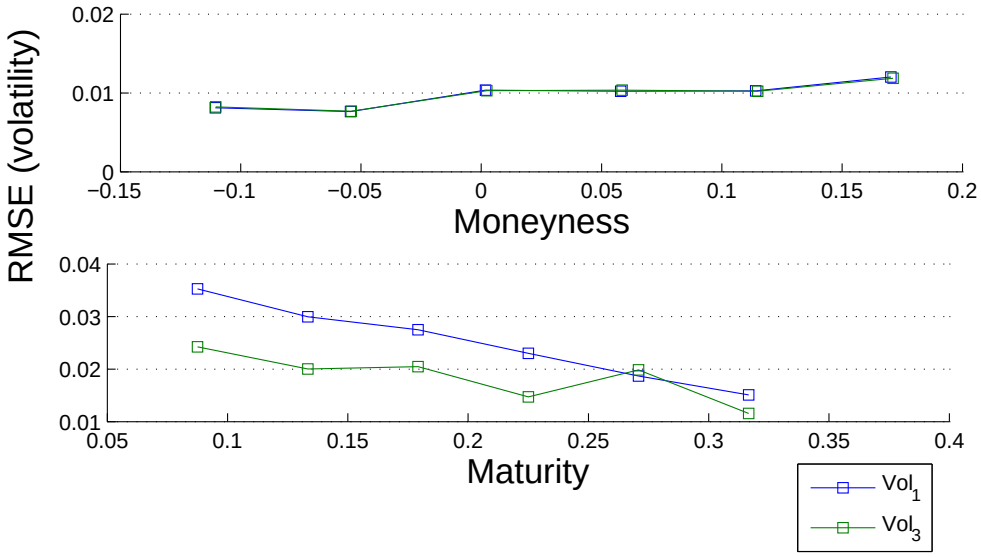

Figure 4.1: Differences in implied volatility between market and model in the case of Vol_3. Here, the RMSRE error metric was used with $\lambda = 3$.

Table 4.8: Residual errors for out-of-sample fitting of Vol_1 and Vol_3. The model was tested against option price data recorded between the months of February and May 2003.

Out-of-sample Errors	Vol_1	Vol_3
$RMSE$	0.842	0.925
$RMSE^{original}$	0.00149	0.00161
$RMSRE(\%)$	3.81%	3.86%
MAE	\$1.40	\$1.64
$MRAE$	0.0203	0.0210

4.8 Discussion

In this chapter we have introduced the second order offset function, derived the corresponding p.d.e. which we have numerically solved using an extension for the finite difference scheme described in Chapter 3. We test a range of volatility specifications, and optimise the associated parameter set of each along with λ. We optimise with respect to the adjusted dataset in the untransformed space, and all errors are measured with respect to the unadjusted prices. We find a significant increase in model fit to market data. Comparing with the original model in Chapter 3

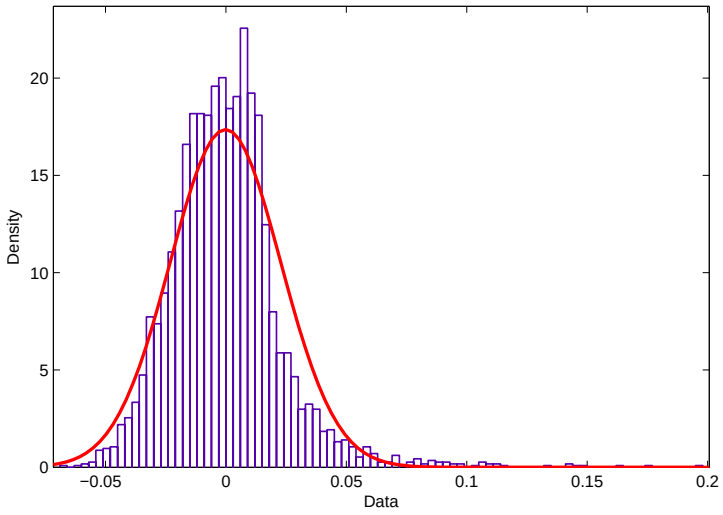

Figure 4.2: Distribution of differences between market implied volatility and model implied volatility in the case of Vol_1, with a fitted normal density function. See Table 4.5 for the mean and standard deviation of this fit.

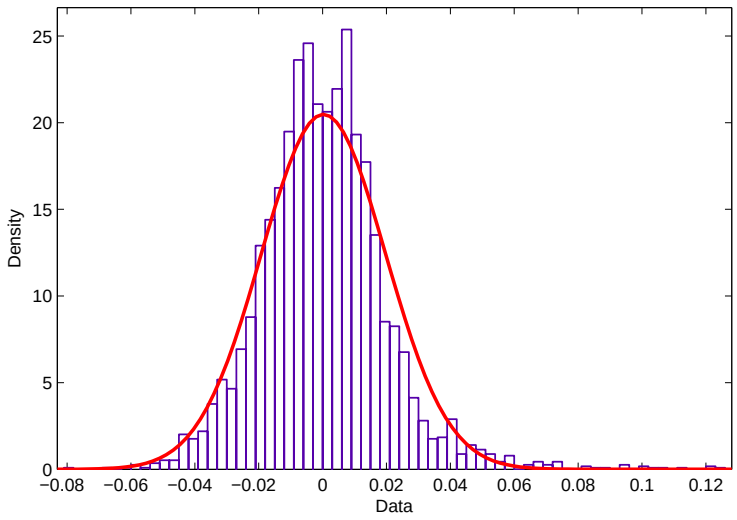

Figure 4.3: Distribution of differences between market implied volatility and model implied volatility in the case of Vol_3, with a fitted normal density function. See Table 4.6 for the mean and standard deviation of this fit.

where RMSRE was found to be of the order of 13%, the model and fitting method presented in this chapter reduces the RMSRE to the order of 3%. Neither do we see any major loss in quality of fit when fitting against an out-of-sample dataset. We can conclude from these results that the inclusion of $D_t^{(2)}$, along with the data adjustment technique, whereby we make use of a daily adjustment index to modify the dataset pre-calibration, and adjust back afterwards, that significant modelling gains can be

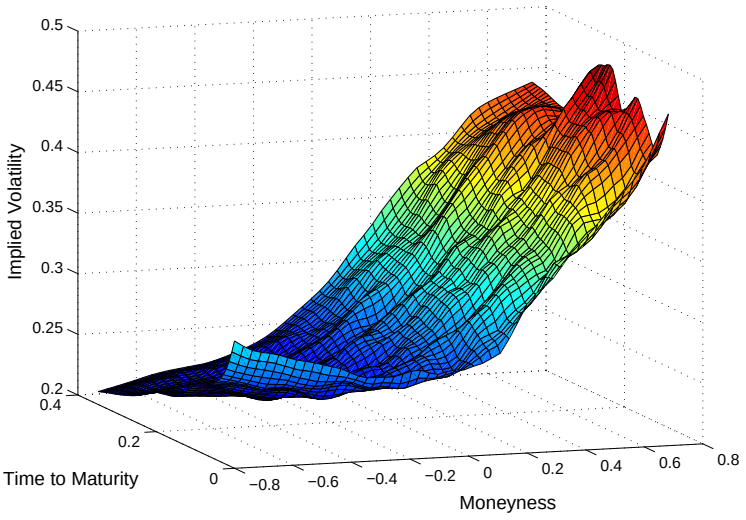

Figure 4.4: Volatility smile for Vol_3.

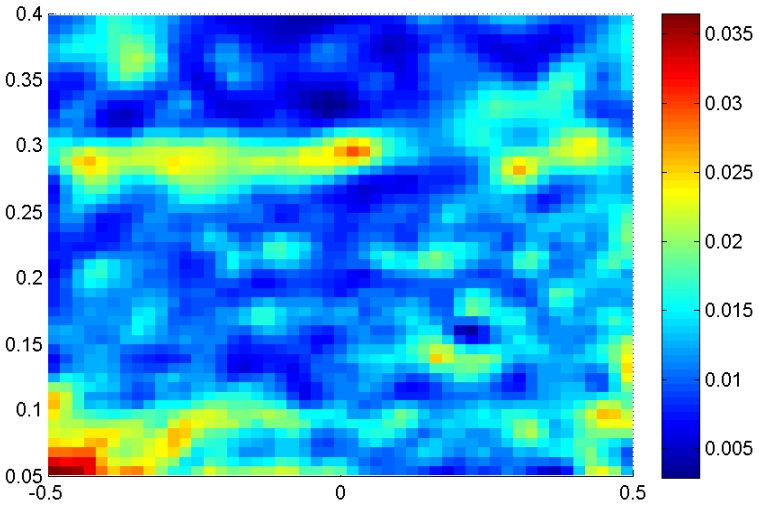

Figure 4.5: Differences in implied volatility between market and model in the case of Vol_3.

made. We have also seen from a quick analysis of the distribution of the errors, that there is no systematic overestimation of implied volatilities that we found in Chapter 3. This can be seen from Figures 4.2 and 4.3. Another point to consider is the speed of the calibration. While we see much better results, the time taken to calibrate this model is an order of magnitude longer than when calibrating any of the models in Chapter 3. This will impact the model's usefulness. Ideally we would like to minimise the number of model parameters while maintaining the accuracy of the model.

Chapter 5

Further results, Summary and Future Work

5.1 Introduction

In this chapter we aim to tackle a number of smaller topics related to the previous work. The first area we cover is the idea of a generalised offset function. In the standard offset function, past information is weighted exponentially, with the greatest emphasis being on more recent data. We study a more generalised offset function in which the past information is discounted at an arbitrary rate. Secondly, we adapt Gatheral's approach [16] to the Heston model in an attempt to derive a semi-analytical closed form solution to the Hobson and Rogers p.d.e. While this approach is not successful, it may point the way towards a solution. Thirdly, we look more closely at actual option prices returned by the model. We note in particular that the model is less effective for large option strike values. Finally, we look at the specific effect of each parameter on the implied volatility smile, essentially calculating the first order sensitivities of the volatility smile to a change in these parameters. Such an analysis is intended to determine if particular parameters are responsible for specific features of the surface. This knowledge would allow users of the model to determine

how to adjust the model to specific market conditions. We finish this chapter with a summary of the thesis and recommendations for future work in this area.

5.2 A generalised offset function

We now consider the more general offset function defined by

$$G_t^{(m)} = \frac{1}{R(t)} \int_0^\infty r(u)(Z_t - Z_{t-u})^m du$$

where $r(t) > 0$ and integrable on $[0, \infty)$ and

$$\frac{1}{R(t)} \int_0^t r(u)du = 1.$$

This approach is also used in [27].

Proposition 22. *The offset process $G_t^{(m)}$ satisfies the coupled s.d.e.'s*

$$dG_t^{(m)} = mG_t^{(m-1)}dZ_t + \frac{m(m-1)}{2}R(t)G_t^{(m-2)}d\langle Z\rangle_t - \frac{r(t)}{R(t)}G_t.$$

Proof. Let $s = t - u$. Then

$$\begin{aligned} R(t)G_t^{(m)} &= -\int_t^{-\infty} r(s)(Z_t - Z_s)^m ds \\ &= \sum_{i=0}^{m} \binom{m}{i} (Z_t)^i \int_{-\infty}^t r(s)(-Z_s)^{m-i} ds. \end{aligned}$$

Now,

$$\begin{aligned} d(R(t)G_t^{(m)}) &= \sum_{i=0}^{m}\binom{m}{i}\Big\{\Big(\int_{-\infty}^{t} r(s)(-Z_s)^{m-i}ds\Big)\times\Big(i(Z_t)^{i-1}dZ_t+\frac{i(i-1)}{2}(Z_t)^{i-2}d\langle Z\rangle_t\Big) \\ &\quad +(Z_t)^i r(t)(-Z_t)^{m-i}dt\Big\} \\ &= m\sum_{i=1}^{m}\binom{m-1}{i-1}\Big(\int_{-\infty}^{t} r(s)(-Z_s)^{(m-1)-(i-1)}ds\Big)(Z_t)^{i-1}dZ_t \\ &\quad +\frac{m(m-1)}{2}\sum_{i=2}^{m}\binom{m-2}{i-2}\Big(\int_{-\infty}^{t} r(s)(-Z_s)^{(m-2)-(i-2)}ds\Big)(Z_t)^{i-2}d\langle Z\rangle_t \\ &\quad +r(t)\sum_{i=0}^{m}\binom{m}{i}(Z_t)^m(-1)^{m-i}dt \end{aligned} \tag{5.1}$$

Note that

$$\begin{aligned} \sum_{i=0}^{m}\binom{m}{i}(Z_t)^m(-1)^{m-i} &= (Z_t)^m\sum_{i=0}^{m}\binom{m}{i}(-1)^{m-i} \\ &= (Z_t)^m(1-1)^m \\ &= 0 \end{aligned}$$

We may rewrite (5.1) as

$$G_t r(t)dt + R(t)dG_t^{(m)} = mR(t)G_t^{(m-1)}dZ_t + \frac{m(m-1)}{2}R(t)G_t^{(m-2)}d\langle Z\rangle_t$$

and so

$$dG_t^{(m)} = mG_t^{(m-1)}dZ_t + \frac{m(m-1)}{2}R(t)G_t^{(m-2)}d\langle Z\rangle_t - \frac{r(t)}{R(t)}G_t.$$

□

In the case of $m = 1$ we have

$$dG_t = dZ_t - \frac{r(t)}{R(t)}G_t dt.$$

Proposition 23. *The price of a European call option, denoted by $f = f(P, G, T-t)$,*

obeys the following partial differential equation,

$$-f_t + rPf_P + -\frac{r(t)}{R(t)}G_t f_G + \frac{\sigma^2}{2}\bigg(P^2 f_{PP} + f_{GG} - f_G + 2\sigma^2 P f_{PG}\bigg) = 0,$$

with boundary conditions $f(P_T, G_T, 0) = \max(P_T - K, 0)$.

Proof. Applying the same change of measure as used in Section 2.1, $dW^{\mathbb{P}} = \phi + dW^{\mathbb{Q}}$, where $\phi = -(\mu/\sigma) - \frac{1}{2}\sigma$, we may write the s.d.e. for our underlying variables as

$$\begin{aligned} dG_t &= (\mu - \frac{r(t)}{R(t)}G_t)dt + \sigma(\phi + dW^{\mathbb{Q}}) \\ &= -\Big(\frac{r(t)}{R(t)}G_t + \frac{\sigma^2}{2}\Big)dt + \sigma dW^{\mathbb{Q}} \qquad (5.2) \\ dP_t &= rP_t dt + \sigma P_t dW^{\mathbb{Q}}. \qquad (5.3) \end{aligned}$$

With this we can now write down a p.d.e. for $f = f(P_t, G_t, T - t)$. We now drop the subscript t, denoting a process, so that its use from here will denote a partial derivative with respect to time. We have by the Feynman-Kac theorem, using equations (5.2) and (5.3), that

$$-f_t + rPf_P - \frac{r(t)}{R(t)}G_t f_G + \frac{\sigma^2}{2}\bigg(P^2 f_{PP} + f_{GG} - f_G + 2\sigma^2 P f_{PG}\bigg) = 0,$$

with boundary conditions $f(P_T, G_T, 0) = \max(P_T - K, 0)$. □

By the same transformation used in Section 2.2 we have

$$\frac{\sigma(x-y)^2}{2}\Big(V_{xx} - V_x\Big) - \frac{r(T-\tau)}{R(T-\tau)}V_y - \lambda V_\tau = 0, \qquad (5.4)$$

where we have used

$$\begin{aligned} x &= \log(\frac{P}{K}) - r(T-t), \\ y &= \log(\frac{P}{K}) - r(T-t) - G, \\ \tau &= -(T-t), \end{aligned}$$

and

$$f(P, G, T-t) \rightarrow Ke^{r(T-t)}V(x, y, \tau).$$

Note that (5.4) belongs to the subclass of Hormander pde's today known commonly as Kolmogorov or Ornstein-Uhlenbeck type [28][29]. In particular the pde (5.4) with boundary condition $V(x, y, 0) = (e^x - 1)^+$ is studied in [30], where conditions are given for the existence and uniqueness of solutions.

Assuming that G may be easily evaluated, the solution of equation (5.4) poses no extra numerical difficulty in the finite difference scheme. To implement the above scheme the only change needed from the scheme in Chapter 3 is the once off pre-evaluation of the integral G at every point on the finite difference grid. This relatively simple extension will allow an examination of different forms of $r(t)$. Currently, with $r(t) = e^{-\lambda t}$, historic log 'returns' Z_t/Z_{t-i} are (negatively) exponentially weighted for increasing i. This approach is analogous to the GARCH discrete model as described in Chapter 2, but unlike GARCH, assumes continuously decreasing weights. The generalisation described here will allow for arbitrary weights to be used. In fact we may consider any functional form of $r(t)$ that satisfies

$$\frac{1}{R(t)} \int_0^t r(u)du = 1.$$

5.3 Fourier transform approach to solution of the Hobson and Rogers p.d.e.

We now attempt to solve the Hobson and Rogers p.d.e. using Fourier transform methods. We take an approach parallel to the derivation of the Heston semi-analytical closed form solution as shown in Chapter 2. We firstly take the original Hobson and Rogers p.d.e. and apply a slightly different change of variables. We use the mapping

$$f(P, D, t) \rightarrow e^{r(T-t)}KV(x, y, \tau).$$

where

$$\begin{aligned} x &= \log(P/K) - r(T-t), \\ y &= D, \\ \tau &= -\lambda(T-t) \end{aligned}$$

Under this transformation of variables our p.d.e. becomes

$$-V_\tau + yV_y - \frac{\sigma(y)^2}{2\lambda}(V_y + V_{xx} - V_x + 2V_{xy} + V_{yy}) = 0.$$

Now, keeping in line with the Heston approach, we assume that the solution of this equation is of the form

$$V(x,y,\tau) = K\big(e^x \Pi_1(x,y,\tau) - \Pi_0(x,y,\tau)\big).$$

Using this as a trial solution we substitute it into the above p.d.e., resulting in the following p.d.e.s for Π_1 and Π_2,

$$-\frac{\sigma(y)^2}{2\lambda}\left(\partial_x\Pi_1 + 3\partial_y\Pi_1 + \partial_{yy}\Pi_1 + 2\partial_{xy}\Pi_1 + \partial_{xx}\Pi_1\right) + y\partial_y\Pi_1 + \partial_\tau\Pi_1 = 0,$$

and

$$-\frac{\sigma(y)^2}{2\lambda}\left(\partial_y\Pi_0 + \partial_{yy}\Pi_0 - \partial_x\Pi_x + 2\partial_{xy}\Pi_0 + \partial_{xx}\Pi_0\right) + y\partial_y\Pi_0 + \partial_\tau\Pi_0 = 0.$$

These equations may be combined, with $j = 0, 1$, as

$$\begin{aligned} -\frac{\sigma(y)^2}{2\lambda}\left((2j+1)\partial_y\Pi_j + \partial_{yy}\Pi_j + (2j-1)\partial_x\Pi_j + 2\partial_{xy}\Pi_j + \partial_{xx}\Pi_j\right) & \\ + y\partial_y\Pi_j + \partial_\tau\Pi_j &= 0. \end{aligned}$$

We now apply the Fourier transform in the x-direction to this p.d.e. The Fourier transform is defined by

$$\tilde{\Pi}_j(u,y,\tau) = \int_0^\infty e^{-iux}\Pi_j(x,y,\tau)dx.$$

Transforming the above p.d.e. we now have

$$\begin{aligned} -\frac{\sigma(y)^2}{2\lambda}\Big((2j+1)\partial_y\tilde{\Pi}_j + \partial_{yy}\tilde{\Pi}_j + (2j-1)iu\tilde{\Pi}_j + 2iu\partial_y\tilde{\Pi}_j - u^2\tilde{\Pi}_j\Big) & \\ +y\partial_y\tilde{\Pi}_j + \partial_\tau\tilde{\Pi}_j &= 0, \end{aligned}$$

which we may rewrite as

$$-\frac{\sigma(y)^2}{2\lambda}\Big(\alpha\tilde{\Pi}_j + \beta\partial_y\tilde{\Pi}_j + \partial_{yy}\tilde{\Pi}_j\Big) + y\partial_y\tilde{\Pi}_j + \partial_\tau\tilde{\Pi}_j = 0. \tag{5.5}$$

where

$$\begin{aligned} \alpha &= -u^2 + (2j-1)iu, \\ \beta &= (2j+1) + 2iu. \end{aligned}$$

The main difference between this p.d.e. and that of Heston is the presence of the $\sigma(y)^2$ term whose functional form is unspecified. To find a general solution, independently of σ we will need to find a solution $\tilde{\Pi}_j$ that solves

$$\alpha\tilde{\Pi}_j + \beta\partial_y\tilde{\Pi}_j + \partial_{yy}\tilde{\Pi}_j = 0. \tag{5.6}$$

As a trial solution we set $\tilde{\Pi}_j = e^{\Lambda y}R(u,\tau)$. This gives us

$$\alpha\tilde{\Pi}_j + \beta\partial_y\tilde{\Pi}_j + \partial_{yy}\tilde{\Pi}_j = e^{\Lambda y}R(u,\tau)(\alpha + \beta\Lambda + \Lambda^2).$$

The choice of

$$\Lambda_\pm = \frac{-\beta \pm \sqrt{\beta^2 - 4\alpha}}{2}$$

ensures that the expression multiplying our $\sigma(y)^2$ term evaluates to zero. We have

$$\tilde{\Pi}_j = (A\exp\{\Lambda_+ y\} + B\exp\{\Lambda_- y\})R(u,\tau).$$

where A and B are arbitrary constants. We may now consider variations of our trial solution such that equation (5.6) holds true, while solving the second part of our p.d.e., namely

$$y\partial_y\tilde{\Pi}_j + \partial_\tau\tilde{\Pi}_j = 0. \tag{5.7}$$

Continuing with the trial solution derived so far, we see that remaining part of our p.d.e. can be written as

$$\begin{aligned} y\partial_y\tilde{\Pi}_j + \partial_\tau\tilde{\Pi}_j &= R(u,\tau)y\partial_y(A\exp\{\Lambda_+ y\} + B\exp\{\Lambda_- y\}) \\ &\quad + A\exp\{\Lambda_+ y\} + B\exp\{\Lambda_- y\}\partial_\tau(R(u,\tau)) \\ &= R(u,\tau)y(A\Lambda_+\tilde{\Pi}_j + B\Lambda_-\tilde{\Pi}_j) \\ &\quad + A\exp\{\Lambda_+ y\} + B\exp\{\Lambda_- y\}\partial_\tau(R(u,\tau)) \end{aligned} \tag{5.8}$$

We need to find a solution such that (5.7) evaluates to zero. We may choose $R(u,\tau) = C(u)e^{-g\tau}$ where g is an arbitrary constant. Then, using equation (5.8), we require

$$\begin{aligned} R(u,\tau)y(A\Lambda_+\tilde{\Pi}_j + B\Lambda_-\tilde{\Pi}_j) - gR(u,\tau)(A\exp\{\Lambda_+ y\} + B\exp\{\Lambda_- y\}) &= 0 \\ \implies R(u,\tau)y(A\Lambda_+\tilde{\Pi}_j + B\Lambda_-\tilde{\Pi}_j) - g\tilde{\Pi}_j &= 0 \\ \implies C(u)e^{-g\tau}y(A\Lambda_+ + B\Lambda_-) - g &= 0 \end{aligned}$$

Since our equation has a y-dependence, we can only have a solution if $A = -(B\Lambda_-)/\Lambda_+$ and $g = 0$ thus no dependence of our solution on time to maturity. Clearly isn't appropriate and so forces us to abandon the approach. This approach in fact shows it is not possible to find a solution by eliminating the coefficient of $\sigma(y)^2$ as the first step. Another possible approach is to initially find a solution such that the expression in equation (5.7) evaluates to a function of the form $(\sigma(y)^2/2\lambda)R(u,y,\tau)$,

again allowing cancellation of the $\sigma(y)^2$ term. The non-linearity of our p.d.e makes it difficult to find such a function, but a careful choice of the functional form of $\sigma(y)$ may be key in finding an overall solution.

5.4 A further analysis of model vs. option prices

In our calibrations, we removed all option prices of value less than \$10. We did this firstly to remove percentage bias when calculating percentage errors. For small option prices, the percentage error can be large even while the absolute error remains low. This can be seen clearly in Figure 5.1. Secondly, the inclusion of options that were deep out of the money resulted in larger errors in model fit. A conclusion can be drawn that the models presented here do not reproduce well the value of out-of-the-money options. We can see this directly by looking at actual option data as a function of strike only. Figure 5.2 presents absolute percentage errors for a selection of option prices from the full market data set. We see a large variability in the model error. On closer examination, (bottom left panel) we see that the errors grow and drop off periodically. It turns out that the dataset is arranged such that option prices, quoted simultaneously, with equal time to maturity, are listed in order of increasing strike. See Table 5.1 below for a sample set of data. The panel in Figure 5.2 shows us that model error grows with increasing strike i.e. as we move further out-of-the-money. The strikes in our dataset range from \$600 to \$900. The error here, for large strikes, may reflect the inability of our model to capture prices in this region, or may reflect noise due to the lack of liquidity of deep out-of-the-money options, and possibly other factors. See Table 5.2. In general, it may be a good strategy to only calibrate to liquid options, or if including a full range of strikes in the calibration, to place a greater weight, in the calibration routine, on those options of higher liquidity. Another solution is to use put-call parity to imply call option prices from put option prices of equal strike and maturity, for low strike values. Since puts will be more

Table 5.1: Model vs Market option prices with absolute relative error as defined by $(f-\hat{f})/\hat{f}$, where as usual f denotes model prices and $\hat{f}$ denotes market prices. The graphical representation of this data can be seen in Figure 5.2. We observe that the error increases for increasing strike. This data represents two sets of option data, quoted simultaneously, but with different times to maturity.

Time to Maturity	Strike	Market Price	Model Price	Absolute Percentage Relative Error
0.171	600	\$224.229	\$224.01	0.0982%
0.171	625	\$200.15	\$199.84	0.155%
0.171	650	\$176.52	\$175.89	0.3539%
0.171	675	\$153.34	\$152.32	0.6684%
0.171	700	\$130.73	\$129.33	1.0757%
0.171	725	\$109.09	\$107.1	1.8241%
0.171	750	\$88.44	\$86.01	2.746%
0.171	775	\$69.32	\$66.48	4.0937%
0.171	800	\$52.16	\$49.	6.067%
0.171	825	\$37.25	\$34.16	8.2854%
0.171	850	\$25.02	\$22.32	10.7558%
0.171	875	\$15.69	\$13.58	13.4783%
0.171	900	\$9.28	\$7.7	17.0377%
0.345	600	\$225.95	\$226.52	0.2543%
0.345	650	\$180.22	\$181.05	0.4617%
0.345	675	\$158.25	\$158.83	0.3625%
0.345	700	\$136.97	\$137.29	0.233%
0.345	725	\$116.58	\$116.68	0.0855%
0.345	750	\$97.31	\$97.2	0.1154%
0.345	775	\$79.48	\$79.11	0.4742%
0.345	800	\$63.14	\$62.69	0.7244%
0.345	825	\$48.72	\$48.17	1.1388%
0.345	850	\$36.36	\$35.74	1.6955%
0.345	875	\$26.16	\$25.55	2.3275%
0.345	900	\$18.02	\$17.52	2.77%
0.345	925	\$11.96	\$11.53	3.5532%
0.345	950	\$7.65	\$7.26	5.0778%

liquid in this region, this strategy may result in a better model fit.

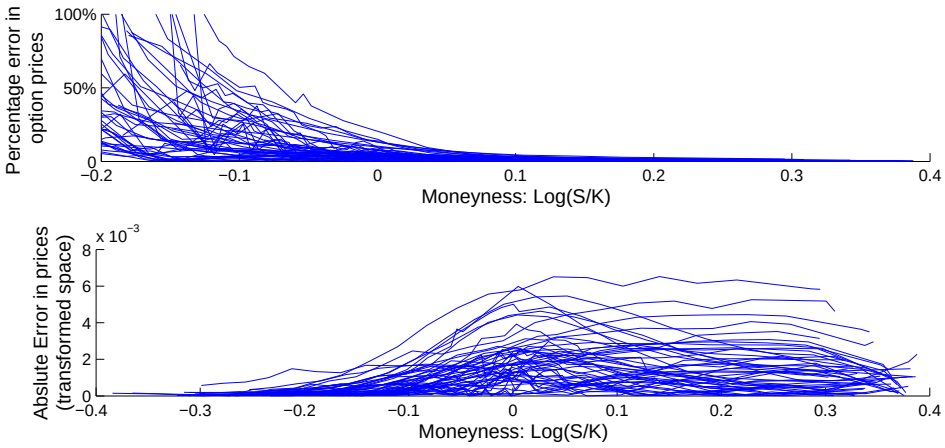

Figure 5.1: A comparison of percentage errors versus absolute errors, plotted as a function of moneyness. The model used here was Vol_3, with the RMSRE calibrated parameters. Each line in the above plots represents a set of options, quoted simultaneously, with the same time to maturity, but over a range of strikes.

5.5 Impact of volatility parameters on implied volatility surface

In models such as the SABR model, or the Heston model, we can attribute the model parameters to specific features relating to market dynamics. For example, the SABR model [12], is given by

$$\begin{aligned} dF_t &= \sigma_t F_t^{\beta} dW_t, \\ d\sigma_t &= \alpha \sigma_t dZ_t, \end{aligned}$$

where W_t and Z_t are two correlated Wiener processes with correlation coefficient $-1 < \rho < 1$, and F_t is a forward price. Here, the α parameter directly corresponds to the volatility of volatility. The β parameter controls the relationship between volatility and price, in that $\beta < 1$ implies an inverse relationship between forward price and volatility. The value of β will determine the 'skewness' of the implied volatility surface. Finally, ρ is the correlation between the forward price process and volatility process. Within the models presented here, we would like to determine if we

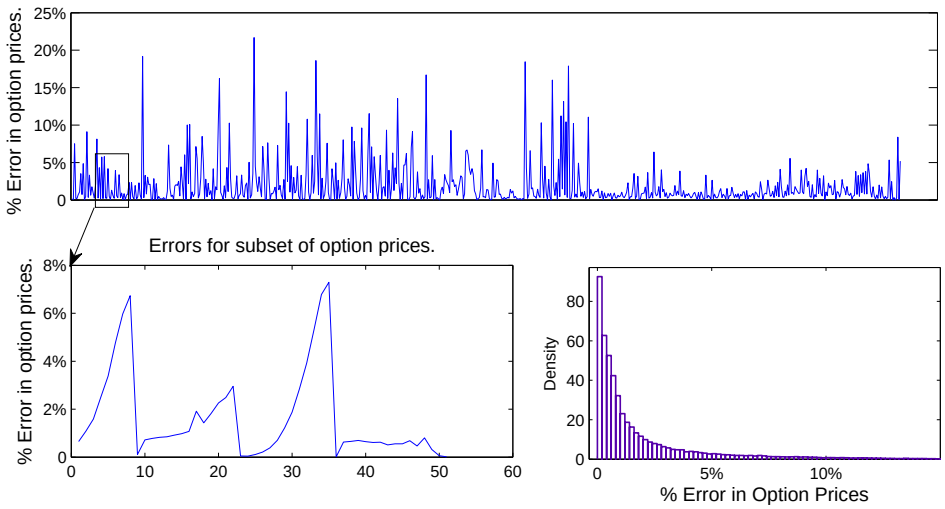

Figure 5.2: Error analysis for the calibrated Vol_3 model. The top panel shows the absolute percentage error from a selection of option quotes picked randomly from the full data set. The data is ordered such that options quoted simultaneously, of equal maturity, and increasing strike are placed beside each other. The panel on the bottom left shows a close-up. On the bottom right panel we see the distribution of the absolute percentage errors.

can attribute any of the parameters of our volatility specification to particular features of the volatility smile. If it were possible to do this, it may lead to a more intuitive and quicker calibration procedure. It would also help in understanding the degree of flexibility the model has to fitting various volatility surface structures. Ideally, we would like to be able to associate three parameters to control the smile, skew, and overall volatility level, combined with another parameter which might describe the volatility term structure. To perform this analysis, we have taken the Vol_1 model, namely

$$\sigma(D_t^{(1)}, D_t^{(2)}) = \alpha_1 + \alpha_2(D_t^{(1)} - \alpha_3)^2 + \alpha_4 D_t^{(2)},$$

from Chapter 4, and calculated the implied volatility smiles for options with fixed maturity equal to 0.1, over a range of strikes. In the calculation of the implied volatilities, we have taken the calibrated parameters, and applied a shift of $\pm 10\%$ to each of the four parameters individually. From the resulting model prices, we calculated the implied volatility smiles which we plotted as a function of moneyness.

Table 5.2: Table of call prices from our market data set versus the implied call prices from put quotes. It is possible, due to lack of liquidity or the impact of transaction costs, that call prices for deep out of the money options may not reflect the true value of those options. In that case calibration to these options prices would lead to a poor model fit. Here we show that put-call parity does not hold in our market data set. A better calibration strategy may be, in the case of out of the money call options, to calibrate to the put-implied call option values calculated via put-call parity.

call price	put-implied call price
\$271.8	\$273.7
\$246.8	\$248.8
\$222.1	\$224.
\$197.8	\$199.5
\$173.5	\$175.
\$149.4	\$151.1
\$139.9	\$141.6
\$130.5	\$132.2
\$125.8	\$127.4
\$121.2	\$122.9
\$112.	\$113.6
\$102.7	\$104.5
\$94.	\$95.7
\$85.2	\$86.9
\$80.7	\$82.7
\$76.7	\$78.2

Essentially, we are calculating the first order sensitivities of the volatility smile to each of the volatility parameters. These smile sensitivity plots are presented in Figure 5.3. We don't expect to see any change in the case of α_1. This is due to the calibrated α_1 parameter being very close to zero, and so a 10% shift will not alter the model prices. The parameters α_2 and α_3 have the greatest affect on the volatility smile. The α_2 parameter corresponds to the scaling constant being applied to the shifted offset function value in our model, while the α_3 specifies the shift. These two parameters appear to control the overall level of the implied volatility surface. Finally α_4 has an affect that can be noticed only for deep in-the-money and out-of-the-money options. This might possibly indicate a relationship between $D_t^{(2)}$ and the skewness of the distribution of returns. We also examined the dependence of the implied volatility as a function of time to maturity, for fixed strike, which resulted in the same conclusions

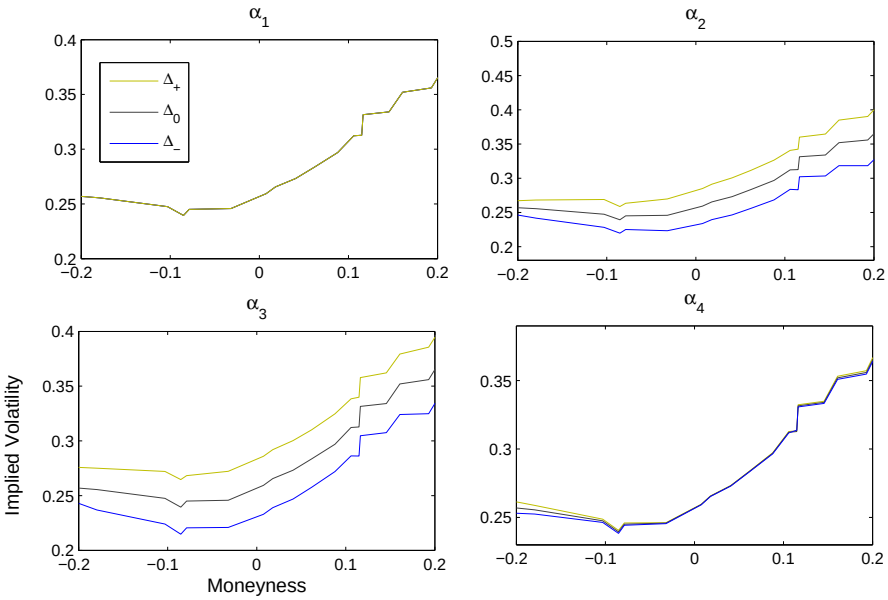

Figure 5.3: Sensitivity of the implied volatility surface to shifts in parameters of the Vol_3 model. Δ_0 corresponds to the unshifted parameters.

as above. This means it is not possible to associate any of the parameters of the model with a specific term structure.

5.6 Summary and Future Work

In this section we provide a summary of the thesis, conclusions that may be reached, and point out a number of directions for future work in this area.

We started in Chapter 1 by presenting a detailed analysis of the derivation of the Black-Scholes p.d.e. We found the original portfolio derivation, as presented in standard text books, to be lacking in some respects. We presented a more rigorous self-financing condition, and demonstrated how a risk free, self-financing portfolio may be used in order to derive the Black-Scholes p.d.e. using a portfolio argument. The derivation presented overcomes the shortcomings in the original derivation. For completeness, and to form a basis for work in later Chapters, we also demonstrate the derivation of the Black-Scholes p.d.e. via a martingale approach, and present a worked solution.

Chapter 2 sees the introduction of the Hobson and Rogers model. We present the original derivation of the model. We also discretise the model which allows a direct comparison with the GARCH model. The direct comparison is useful to indicate the typical parameters we should expect to obtain in the calibration of the Hobson and Rogers model. A significant difference between the two modelling approaches is that while GARCH considers daily returns, the Hobson and Rogers model takes as input, returns defined by the ratio of today's price, to some historic price not necessarily quoted on the preceding day. Despite this, we demonstrate how to overcome this difficulty to make a direct comparison between these two models.

Chapter 3 follows the approach used in [22] to implement a numerical scheme to solve the Hobson and Rogers p.d.e. under transformed variables. We attempt to reproduce the results from [22] and we also explore a number of other functional forms of the volatility specification. Firstly, we show that the choice of λ, the parameter which controls the weighting of past data in the offset function, is an important parameter in fitting the model to market data. We found that the choice of λ depends on the volatility specification being used. We also consider the choice of error metric used in the calibration. The choice of metric is important as it places more emphasis on in-the-money options in the case of relative error, and out-of-the-money options in the case of absolute error. We also note that it is important to minimise the residual using the untransformed option price. Analysis of the market data and the dependence of the volatility surface on time has shown that calendar time has a significant impact on implied volatility levels. Since the volatility specification has no time dependence, it will not be able to capture this calendar time dependence. This will result in a poorer model fit. To counteract this problem we implement a scheme to modify the market data used in our model by use of a daily adjustment index. We calibrate to adjusted data, and then readjust back. Overall we see a significant improvement in model fit.

In Chapter 4 we extend the volatility specification to now depend on the second order offset function. Our option price now depends on both the first and second order offset functions, along with underlying price and time to maturity. We derive the

corresponding p.d.e. and apply a very useful transformation of variables which allows us to easily implement a numerical scheme. The finite difference scheme from Chapter 3 is adapted to include the extra dimension of $D_t^{(2)}$. Again we explore a number of volatility specifications which include $D_t^{(2)}$ and again see a significant improvement in model fit across all error metrics.

Finally, in Chapter 5, we present some additional calculations and analysis. We have derived the p.d.e. associated with a generalised offset function and from this result it is clear that, from a numerical point of view, it would not be difficult to implement this generalisation. We attempt to solve the Hobson and Rogers p.d.e. in a similar fashion to the derivation of the solution to the Heston model. We make some further analysis of model prices, and discuss the possibility of errors that may arise due to pricing out-of-the money options. We also conduct a first order sensitivity analysis of the model containing the Vol_3 volatility specification.

Suggestions for further work in this area would be to numerically explore the more generalised offset functions, and to consider alternative functional forms of $\sigma(D_t^1, \ldots, D_t^n)$. A semi-analytical closed form solution, if found, would offer a significant increase in usability by making this model very quick to calibrate. By careful choice of $\sigma(D_t^1, \ldots, D_t^n)$ it should be possible to find such a solution.

Bibliography

[1] D G Hobson and Rogers. L.C.G.: Complete models with stochastic volatility. *Math. Finance*, (8), 1998.

[2] Black, F. and Scholes, M. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3), 1973.

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

[4] Mikosch, T. *Elementary stochastic calculus with finance in view*, volume 6. World Scientific Pub Co Inc, 1998.

[5] Björk, T. *Arbitrage theory in continuous time.* Oxford University Press, USA, 2004.

[6] Andrea Pascucci. *PDE and martingale methods in option pricing*, volume 2. Springer, 2011.

[7] B.K. Øksendal. *Stochastic differential equations: an introduction with applications.*

[8] Carr, P.P. and Bandyopadhyay, A. How to Derive Black-Scholes Equation Correctly? *Urbana*, 51:61801.

[9] Davis, M.H.A. Black–Scholes Formula. *Encyclopedia of Quantitative Finance*, 2010.

[10] Muller, P. H. E. B. Dynkin, Markov Processes. Volume I und II. (Die Grundlehren der mathematischen Wissenschaften, Band 121/122) VII + 365 S.

und VIII + 274 S. Berlin/Heidelberg/New York 1965. Springer-Verlag. Preis geb. DM 96. *ZAMM - Journal of Applied Mathematics and Mechanics / Zeitschrift für Angewandte Mathematik und Mechanik*, 46(1):74–74, 1966. ISSN 1521-4001. doi: {10.1002/zamm.19660460123}. URL `http://dx.doi.org/10.1002/zamm.19660460123`.

[11] Affleck-Graves, J. and McDonald, B. Nonnormalities and tests of asset pricing theories. *Journal of Finance*, pages 889–908, 1989.

[12] Patrick S Hagan, Deep Kumar, Andrew S Lesniewski, and Diana E Woodward. Managing smile risk. *Wilmott Magazine*, page 249, 2002.

[13] Rubinstein, M. Implied binomial trees. *Journal of finance*, pages 771–818, 1994.

[14] Poon, S.H. and Granger, C.W.J. Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 41(2):478–539, 2003.

[15] Christian Gouriéroux and C Gourieroux. *ARCH models and financial applications*, volume 1. Springer, 1997.

[16] Gatheral, J. *The volatility surface: a practitioner's guide*, volume 357. Wiley, 2006.

[17] Duffie, D. and Pan, J. and Singleton, K. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.

[18] Scott, L.O. Option pricing when the variance changes randomly: Theory, estimation, and an application. *Journal of Financial and Quantitative analysis*, 22(04):419–438, 2009.

[19] Cox, J.C. and Ross, S.A. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3(1-2):145–166, 1976.

[20] Harrison, J.M. and Pliska, S.R. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11(3): 215–260, 1981.

[21] Karatzas, I. and Shreve, S.E. Brownian motion and stochastic calculus (Graduate Texts in Mathematics 113). *Eds. S. Axler, FW Gehring and PR Halmos. Berlin, Heidelberg: Springer Verlag*, 1988.

[22] Paolo Foschi and Andrea Pascucci. Calibration of a path-dependent volatility model: Empirical tests . *Computational Statistics & Data Analysis* , 53(6):2219 – 2235, 2009. ISSN 0167-9473. doi: http://dx.doi.org/10.1016/j.csda.2008.10.042. URL http://www.sciencedirect.com/science/article/pii/S0167947308005069. The Fourth Special Issue on Computational Econometrics .

[23] Jeantheau, T. A link between complete models with stochastic volatility and ARCH models. *Finance and Stochastics*, 8(1):111–131, 2004.

[24] Hubalek, F. and Teichmann, J. and Tompkins, R. Flexible complete models with stochastic volatility generalising Hobson-Rogers. *preprint*, 2004. URL http://www.fam.tuwien.ac.at/~jteichma/ghr220704.pdf.

[25] Platania, A. Putting the Hobson& Rogers model to the test. *preprint*, 2005. URL http://www.statslab.cam.ac.uk/~chris/papers/AP_CR_HR.pdf.

[26] Marco Di Francesco and Andrea Pascucci. On the complete model with stochastic volatility by hobson and rogers. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 460(2051): 3327–3338, 2004.

[27] Paolo Foschi and Andrea Pascucci. Path dependent volatility. *Decisions in Economics and Finance*, 31(1):13–32, 2008. ISSN 1593-8883. doi: 10.1007/s10203-007-0076-6. URL http://dx.doi.org/10.1007/s10203-007-0076-6.

[28] Lars Hörmander. Hypoelliptic second order differential equations. *Acta Mathematica*, 119(1):147–171, 1967. ISSN 0001-5962. doi: 10.1007/BF02392081. URL http://dx.doi.org/10.1007/BF02392081.

[29] A. Kolmogorov. Zufllige bewegungen (zur theorie der brownschen bewegung). *Ann. of Math.*, II(35):116–117, 1934.

[30] Marco Di Francesco and Andrea Pascucci. On a class of degenerate parabolic equations of kolmogorov type. *Applied Mathematics Research eXpress*, 2005(3): 77–116, 2005.

Appendix A

Matlab Code

Note that the code structure is based on that kindly provided by Paolo Foschi and a proportion of the original code remains unchanged. Details of changes and on the implementation of the 3D finite difference scheme can be found in Sections 3.5.1 and 4.3 respectively.

Listing A.1: Calibration: Sets calibration options, calls the underlying calibration routine and returns the relevant residual error.

```
1 function [p,residuals,jacobian] = calibrate3d(xs,ys1,ys2,
     ts, us, p0,lambda,conversion_factor,error_measure,SI )
2
3 mu = max(abs(xs))+0.01;     I=101;
4 ni = max(abs(ys1))+0.01;    J=81;
5 chi=max(abs(ys2))+0.01;     K=21;
6 T = max(abs(ts))+0.01;      N=15;
7 dx = mu/(I-1)*2;     dy = ni/(J-1)*2;     dz=chi/(K-1)*2;
     dt = T/(N-1);
8 MAX_EVALS = 200;
9
```

```
options = optimset('Jacobian', 'off', ...
    'LevenbergMarquardt', 'on',...
    'Algorithm','interior-point', ...
    'Display','iter',...
    'Diagnostics','on',...
    'LargeScale','off', ...
    'OutputFcn', @outfun, ...
    'MaxFunEvals', MAX_EVALS, ...
    'TolX', 1e-4 );

[p,resnorm,residuals,exitflag,output,lambda,jacobian] =
   ...
    lsqnonlin( @my_eval, p0,[],[], options );

    function [res] = my_eval( p)

        x = ((1:I)-1)*dx-mu;
        y = ((1:J)-1)*dy-ni;
        z = ((1:K)-1)*dz-chi;
        t = -((1:N)-1)*dt;

        U= HR_evaluate3d(p, mu,ni,chi,T, I,J,K,N, ...
            lambda, sigma_max, @payoff_call);

        vs = interpn(x,y,z,t, U, xs,ys1,ys2,ts,'linear' )
           ;

        if error_measure == 1
```

```
36              res = vs-us;
37          elseif error_measure == 2
38              res = (vs-us)./us;
39          elseif error_measure == 3
40              res = (vs-us).*conversion_factor;
41          elseif error_measure == 4
42              res = (((vs-us).*conversion_factor).^2)
                    .^0.25;
43          elseif error_measure == 5
44              res = ((((vs-us).*conversion_factor).^2)
                    .^0.5)./us).^0.5;
45          end
46
47          function u0 = payoff_call(x,y)
48              u0 = max(exp(x)-1,0);
49          end
50
51          function stop =outfun( p, optimValues, state )
52              stop = false;
53          end
54
55      end
56  end
```

Listing A.2: Calibration: Setup of grid.

```
function  U=HR_evaluate3d( p, mu,ni,chi,T, I,J,K,N,
   lambda, sigma_max, ...
    payoff )
dx = mu/(I-1)*2;    dy = ni/(J-1)*2;    dz=chi/(K-1)*2;
   dt = T/(N-1);
[yy,xx,zz] = meshgrid(((1:J)-1)*dy-ni, ((1:I)-1)*dx-mu,
   ((1:K)-1)*dz-chi );
dd1=  xx-yy;
dd2 = (xx-yy).^2-zz./2;
sigma2 = calculate_sigmas(p,dd1,dd2,lambda,sigma_max);
U=newkolmogorov3d( -mu, -ni,-chi, dx,dy,dz,dt, I,J,K,N,
   ...
    sigma2, -sigma2,  @(x,y) (x-y), payoff );
end
```

Listing A.3: Implementation of 3d finite difference scheme

```
function [U] = newkolmogorov3d( x0, y0,z0, dx,dy,dz,dt, I
   ,J,K,N, a, b,c, u0)

[yy,xx,zz] = meshgrid( ((1:J)-1)*dy+y0, ((1:I)-1)*dx+x0 ,
    ((1:K)-1)*dz+z0);
Z = makeZ( I,J, x0,y0, dx,dy,dt, c );
n=I*K;
D1=sparse(1:I*K,1:I*K,0);
D1 = D1+ sparse(2:n,1:n-1,1,n,n) - sparse(1:n,1:n,2) +
   sparse(1:n-1,2:n,1,n,n);
D1 = D1 / (dx*dx);
```

```
D1(1,1)=0;
D1(1,2)=0;
D1(n,n-1)=0;
D1(n,n)=0;

count=I;

while count<=I*(K-1)
    D1(count,count+1) = 0;
    D1(count,count) = 0;
    D1(count,count-1) = 0;
    count=count+1;
    D1(count,count+1) = 0;
    D1(count,count) = 0;
    D1(count,count-1) = 0;
    count=count+I-1;
end

D1=sparse(D1);
D2=sparse(1:I*K,1:I*K,0);
D2 = D2+ sparse(2:n,1:n-1,-1,n,n) + sparse(1:n-1,2:n,1,n,
   n);
D2 = D2 / (2*dx);
D2(1,1)=0;
D2(1,2)=0;
D2(n,n-1)=0;
D2(n,n)=0;
count=I;
```

```
36
37 while count<=I*(K-1)
38     D2(count,count+1) = 0;
39     D2(count,count-1) = 0;
40     count=count+1;
41     D2(count,count+1) = 0;
42     D2(count,count-1) = 0;
43     count=count+I-1;
44 end
45
46 D2=sparse(D2);
47 AA=cell(J,1);
48
49 for j=1:J
50     AA{j}=sparse(1:I*K,1:I*K,0);
51 end
52
53 BB=cell(J,1);
54 for j=1:J
55     BB{j}=sparse(1:I*K,1:I*K,(1+(x0+mod(0:I*K-1,I)*dx-(y0
          +(j-1)*dy)).^2).^0.5);
56 end
57
58 zvalues= sparse(1:I*K,1:I*K,z0 + ceil((1:I*K)/I)*dz );
59
60 for j=1:J
61     da =  squeeze(a(:,j,:));
62     da=sparse(diag(da(:)));
```

```
63      DZ=sparse(I+1:n-I,2*I+1:n,1,n,n)+sparse(I+1:n-I,1:n
            -2*I,-1,n,n);
64      DZ=DZ/(2*dz);
65      AA{j}=sparse(eye(n)-BB{j}*dt*da*(D1-D2-DZ)+zvalues*DZ
            );
66  end
67
68  clear a b c da db D1 D2
69  U = zeros(I,J,K,N);
70  U(:,:,:,1) = u0( xx, yy );
71  vv = U;
72  for n=2:N
73      for k=1:K
74          layer=vv(:,:,k,n-1);
75          vv(:,:,k,n) = reshape(Z*layer(:), I,J );
76      end
77      for j=1:J
78          temp=squeeze(vv(:,j,:,n));
79          vv(:,j,:,n) =reshape( AA{j}\  temp(:),I,1,K);
80      end
81  end
82  U=vv;
```

Listing A.4: Calculation of offset

```
function m = exptrend5( t, z, lambda,power )

lookback = 100;
number_of_elements=size(z);
number_of_elements=number_of_elements(1,1);
m=zeros(number_of_elements,1);
for i=(lookback+1):number_of_elements
    term=zeros(lookback,1);
    count=1;
    weight=zeros(lookback,1);
    while   count <= lookback
        weight(count)=lambda/252*exp(-lambda*(t(i)-t(i-
           lookback+count-1)))*count;
        count=count+1;
    end
    weight=weight/sum(weight);
    count=1;
    while   count < lookback
        j=i-count;
        term(count)=weight(count)*(z(i)-z(j))^power;
        count=count+1;
    end
    m(i,1)=sum(term);
end
```