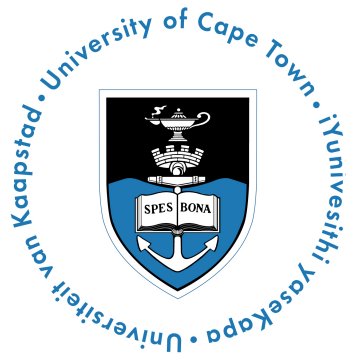


The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.



MPHIL (MATHEMATICAL FINANCE)
DISSERTATION
**Extracting risk aversion estimates from
option prices/implied volatility**

Aveshen Pillay
PLLAVE002
aveshen.pillay@uct.ac.za

18 November 2010

Acknowledgements

I would like to express my gratitude to the following people:

Obeid Mohamed and Greg Judin for allowing me access to a Bloomberg terminal on numerous occasions to download data for this paper.

My supervisor, Shakill Hassan, for his continuous motivation to ensure the success of this paper and to ensure that at no point in time do I accept second best. Your patience and direction throughout this process has made writing this paper a great experience.

To my friends, who have offered their continued support and patience throughout my years at UCT and beyond. I will forever be indebted to you.

Finally, to my parents, for their unwavering support and encouragement throughout my life. You have instilled in me the pursuit of excellence in all aspects of my life. The coming together of this paper is only the first lap in this race. My utmost appreciation to you.

Abstract

The risk neutral density function is the distribution *implied* by the market price of derivative securities, namely options. It encloses the assumption that arbitrage free conditions persist in the market. Given the *historical* evolution of stock prices, an investor will form some belief about the future progression of the stock price. By looking at a historical density function of the return distribution, you can recover this opinion. Firstly, in the absence of market price data of derivative securities, to what extent can other available information be reconciled to yield this risk neutral distribution? Secondly, can we recover the stochastic discount factor and then characteristics of investor risk aversion by specifying a functional form for the pricing kernel and then solving for empirical risk aversion?

We find promising results over the period 2 January 2005 to 31 December 2009 using the S&P500 index as the basis for the study. We are able to fit and calibrate a stochastic volatility model with jumps to the market implied volatility surface and then use this model to extract the risk neutral density and “market” option prices. In addition, we are able to find a time varying pricing kernel. Under a normal and non-parametric historic density we extract relative risk aversion estimates consistent with investor behaviour and the state of the overall economy at the time. Risk aversion is counter cyclical, a result consistent with the work by [Rosenberg and Engle \(2002\)](#) and [Fama and French \(1989\)](#).

Contents

1	Introduction	4
2	The state price density	7
3	Pricing kernels and risk aversion	8
3.1	The empirical pricing kernel	8
3.2	Reconciling pricing methods	10
3.3	Consumption based approach to asset pricing	11
4	Estimation strategy	13
4.1	Specification of the state price density	13
4.1.1	Stochastic volatility models	14
4.1.2	Model calibration	17
4.1.3	Simulating the risk neutral distribution	19
4.2	Specification of the historical density	20
4.2.1	Lognormal density function	20
4.2.2	Empirical innovation density function	20
4.2.3	Nonparametric density function	21
4.3	Estimating the pricing kernel	23
5	Data	25
5.1	S&P500 equity index	25
5.2	S&P500 implied volatility data	26
5.3	S&P500 dividend yield	26
5.4	Risk free interest rates in the United States	26
5.5	VIX index	28
5.6	TED spread	28
6	Results	30
6.1	Summary	30
6.2	Stochastic volatility model calibration	30
6.3	Historical density functions	37
6.3.1	Lognormal density function	37
6.3.2	Empirical innovation density function	37
6.3.3	Nonparametric density function	38
6.3.4	Density function analysis over time	39
6.4	Empirical pricing kernel	42
6.4.1	Empirical risk aversion	43
7	Conclusion	52

A	Mathematical derivations	54
A.1	Deriving the empirical risk aversion parameter	54
A.2	Deriving the distribution of stock prices	54
A.2.1	Ito's formula	54
A.2.2	Derivation	54
B	Mathematical proofs	56
B.1	Risk neutral dynamics of a jump process	56
B.2	Relative entropy	56
C	Option pricing under affine functions	59
C.1	Characteristic function	59
C.2	Fourier transforms	60
C.3	Computing the Fourier integral	60
D	Solving nonlinear least squares problems	61
E	Simulation of probability densities	61
	References	63

University of Cape Town

List of Figures

1	S&P500 Equity Index	25
2	Volatility Surface on 10 June 2009	27
3	Dividend yield 2005 - 2009	27
4	US Treasury Bills Data	28
5	The VIX Index	29
6	The TED spread	29
7	Performance of the Heston Model - 10 June 2009	33
8	Parameters in the Heston Model	34
9	Performance of the Bates Model - 10 June 2009	35
10	Parameters in the Bates Model	36
11	Probability density functions	40
12	Probability plot of returns against Normal and t -distributions	41
13	Empirical pricing kernel in June - Heston	45
14	Empirical pricing kernel in June - Bates	46
15	Empirical price kernel parameters - Heston	47
16	Empirical price kernel parameters - Bates	48
17	Empirical risk aversion	51

List of Tables

1	Empirical characteristics of the S&P500 Equity Index	25
2	Volatility term structure	26
3	Empirical characteristics of the US T-Bill data	28
4	Initial parameter estimate and parameter bounds - Heston model	31
5	Initial parameter estimate and parameter bounds - Bates model	32
6	Estimated parameters in the lognormal density specification	37
7	Estimated parameters in the GARCH density specification	38
8	Estimated parameters in the Non-parametric density specification	39
9	Regression results	50

1 Introduction

In a general equilibrium setting, an Arrow-Debreu security is a contract that agrees to pay one unit at a particular date in the future if a particular state has been realised, and zero in all other states. In a risky environment, these securities are fundamental for the understanding of the time-state preferences of investors. Assuming a multi-period setting and a continuum of states, for every realisation of a state x , the state price density (SPD), which reveals the price of a security that pays one unit if the realised state falls between x and $x+dx$, and zero in all other states, defines the price of an Arrow-Debreu security (Ait-Sahalia and Lo, 1998).

Define a complex security as one that pays off in more than one state realisation. Assuming complete markets, a portfolio of Arrow-Debreu securities can replicate any complex security. To find the price of a derivative (complex) security one can construct a replicating portfolio whereby the cash flows mimic these state contingent claims and use the SPD. In reality, however, markets are incomplete and such state contingent claims are not traded. The question then lies as to whether one can extract characteristics of the SPD, and consequently characteristics of the investor from the market price of traded securities.

There are two routes to find the SPD: preference-based equilibrium theory where the equality of supply and demand yields a price or arbitrage theory in which the fair price of a security is found assuming that there is no positive probability of gaining without a positive probability of also losing. Denote the price obtained under either theory as the model price. Reconciliation of this price to the market price will reveal features of the SPD.

Ait-Sahalia and Lo (1998) non-parametrically recover the SPD from historical option prices. The authors note that in a preference-based equilibrium framework, the SPD can be expressed in terms of a stochastic discount factor (SDF) or pricing kernel such that all discounted asset prices are martingales under the actual distribution of aggregate consumption after multiplication by this SDF. Within a no arbitrage framework the SPD is also known as the risk neutral density and can be used to derive the price of securities assuming all investors are risk neutral (Ross, 1976; Cox and Ross, 1976). Assuming all investors are risk neutral requires that all assets earn an expected return equal to the risk free rate.

Rosenberg and Engle (2002) examine characteristics of an investor over various states of the world by estimating a time varying asset-pricing kernel. They define the pricing kernel as a preference function that reproduces asset prices based on a forecast payoff density (historic density). Although their results are mixed on the correct functional form of the pricing kernel, they find that from a power specification one is able to extract characteristics of investor risk aversion. Furthermore, they suggest that hedge ratios based on a time varying asset pricing kernel improve hedging performance compared to the use of a time

invariant kernel. Their work has inspired a revisit of estimating investor risk aversion over time through the estimation of a time varying pricing kernel. Here however, we suggest a more generalised methodology that only uses market implied volatility data as an input.

In a market where option prices are not readily available or unreliable due to poor liquidity our goal is to test the performance of stochastic volatility models to find whether model prices are consistent with the market prices of traded options. The ability to fit and calibrate a stochastic volatility model not only affords us the capacity to calculate realistic option prices, but also to use such models in the pricing of higher dimensional derivative structures. If we are able to find risk aversion estimates from our time varying pricing kernel that are consistent with the assumption of no arbitrage one may find this methodology useful in emerging markets or illiquid markets. The hedging results from [Rosenberg and Engle \(2002\)](#) for their pricing kernel specification suggest a useful result that can be applied to trading and hedging in emerging or illiquid markets.

Given a volatility input, this paper aims to accomplish two things. Firstly, in the absence of market price data of derivative securities, to what extent can we use other available information to yield the SPD? Secondly, this paper aims to recover the stochastic discount factor and then characteristics of investor risk aversion by specifying a functional form for the pricing kernel and then solving for empirical risk aversion in a similar fashion to [Rosenberg and Engle \(2002\)](#).

Being able to answer these questions offers a useful platform upon which one can understand investor behaviour and the evolution of asset prices in emerging markets or illiquid markets. We find promising results over the period 2 January 2005 to 31 December 2009 using the S&P500 index as the basis for our study. We are able to fit and calibrate a stochastic volatility model to the market implied volatility surface. We find that a stochastic volatility model with jumps ([Bates \(1996\)](#) model) is able to fit the market volatility surface better than the standard stochastic volatility model of [Heston \(1993\)](#). The S&P500 index provides a useful data sample as it is one of the most widely traded indices and the index that is most indicative of market performance in the United States. Our sample period contains the greatest market crash in equity indices since the Great Depression.

We use our stochastic volatility models to extract the risk neutral density and “market” option prices. In addition, we are able to find pricing kernel estimates by estimating a density function from an EGARCH model together with normal and non-parametric density functions of historic returns. The pricing kernel exhibits its time varying nature; from which we are able to extract relative risk aversion estimates. The results from a normal and non-parametric density are consistent with investor behaviour reported in market reports and the state of the overall economy at the time.

The paper progresses as follows: section 2 provides a background on the state price density and how we can go about calculating risk aversion. Section 3 provides insight into the asset-pricing kernel and how we can find this kernel using our density functions; as well as an understanding of this kernel in a consumption based approach to asset pricing. Section 4 extensively covers how we specify our stochastic volatility models and calibration thereof. Our historic density functions are also described. A reader familiar with this subject can jump to section 4.3 where we describe how we estimate our pricing kernel. Section 5 describes our data, while section 6 contains the results. Section 7 concludes.

University of Cape Town

2 The state price density

Given the historical evolution of stock prices, an investor will form some belief about the future progression of the stock price. By looking at a historical density function of the return distribution, one can recover this opinion. Examples include the lognormal distribution of stock prices as assumed in the Black-Scholes world or a stochastic volatility model such as GARCH. Denote this distribution the physical or historical density.

The risk neutral density, however, is the distribution implied by the market price of derivative securities, namely options. This distribution differs from the historical distribution as the risk neutral density uniquely characterises the equivalent martingale measure under which discounted asset prices are martingales (Harrison and Kreps, 1979).

We shall describe the relationship between the probability distributions as defined in Jackwerth (2000):

$$\text{Risk neutral} = \text{Historic} \times \text{Risk aversion adjustment}$$

The risk neutral probability is the price, multiplied by the risk free return that an investor would be willing to pay for receiving one dollar in a particular state. The historic probability distribution is the investors belief of the likelihood of the outcome of a particular state (Jackwerth, 2000).

If an investor were indifferent to risk, then these probabilities would be identical. However, it is known that investors value returns differently in different states of the world and hence the presence of the risk aversion adjustment term. The risk aversion adjustment aims to capture these investor preferences. The goal is then to identify, empirically, characteristics of the risk aversion function given the risk neutral and historic (empirical) distribution.

Historical returns, typically, have been used to estimate probability distributions; however, they may not be able to capture the probability of extreme events, which history has shown, that although rare, do occur.

Historical analysis in Jackwerth and Rubinstein (1996) reveals two well know observations. Namely, that historically measured *volatility* varies significantly over different time intervals and that historical volatility can be a poor predictor of future implied volatility. The response to address this shortcoming is to postulate a particular statistical time series model of *returns*, for example, an ARCH/GARCH model. Other methods include the use of option price data to imply parameters of a pre-specified (risk neutral) stochastic process. We shall use the later approach to find the risk neutral density.

3 Pricing kernels and risk aversion

3.1 The empirical pricing kernel

Having an understanding of Arrow-Debreu securities and the risk neutral density, we can now characterise investor preferences for payoffs over different states of the world using the pricing kernel. A time varying pricing kernel summarises investor preferences for payoffs over different states of the world over time.

Define the payoff of an asset in period t as some function $\Psi(X_t)$ of the asset price X_t . Assuming no arbitrage, current asset price equals the expected pricing kernel - weighted payoff (Rosenberg and Engle, 2002):

$$p_t = \mathbb{E}_t [M_t \Psi(X_t)], \quad (1)$$

where:

p_t	current asset price
M_t	asset pricing kernel
$\Psi(X_t)$	asset payoff in period $t + 1$

Lucas (1978) defines the pricing kernel as function of consumption as follows:

$$M_t = \frac{U'(C_{t+1})}{U'(C_t)}, \quad (2)$$

where:

$U(\cdot)$	representative utility function
C_t	investor consumption in period t

The pricing kernel is characterised by its slope. In addition, risk aversion function parameters are functions of the slope.

Arrow (1971); Pratt (1964) show that relative risk aversion can be calculated as a function of consumption:

$$\gamma_t = - \frac{C_{t+1} \cdot M'_t(C_{t+1})}{M_t(C_{t+1})}, \quad (3)$$

If the pricing kernel is a function of multiple variables, the level of risk aversion will also be a function of those variables. We need to look at ways to estimate the kernel independently of any prior specification of these variables to remove any subjective bias. If the payoff of the asset is $\Psi(X_t)$, our concern is the projection of the pricing kernel onto these payoffs.

Define the original pricing kernel as follows:

$$M_t = M_t(Z_t, Z_{t+1}), \quad (4)$$

where:

Z_t | vector of pricing kernel state variables at time t

Given the subjective nature of the choice of variables Z_t , that would affect the pricing kernel, the projected pricing kernel estimate is postulated as a function of the asset payoff at time $t + 1$.

Now evaluation of the expectation in [1](#) is done in two steps. Recall the result from elementary statistics for density functions, that the joint density function is equal to the product of the marginal and conditional densities.

1. Recover the projected pricing kernel by integrating the pricing kernel using the conditional density:

$$\begin{aligned} M_t^*(X_t) &= \int M_t(Z_t, Z_{t+1}) f_t(Z_{t+1}|X_{t+1}) dZ_{t+1} \\ &= \mathbb{E}_t [M_t(Z_t, Z_{t+1})|X_{t+1}], \end{aligned} \quad (5)$$

Projected pricing kernel $M_t^*(X_t)$ depends on the state of the world next period through X_{t+1} in the same way that the original pricing kernel $M_t(Z_t, Z_{t+1})$ depends on state of the world next period through Z_{t+1} .

2. To recover the asset price p_t , integrate the product of projected pricing kernel and payoff variable with respect to marginal density:

$$\begin{aligned} p_t &= \int M_t^*(X_{t+1}) \Psi(X_t) f_t(X_{t+1}) dX_{t+1} \\ &= \mathbb{E}_t [M_t^*(X_{t+1}) \Psi(X_t)], \end{aligned} \quad (6)$$

Equation [1](#) can now be written as:

$$\begin{aligned} p_t &= \mathbb{E}_t [M_t(Z_t, Z_{t+1}) \Psi(X_t)] \\ &= \mathbb{E}_t [\mathbb{E}_t (M_t|X_{t+1}) \Psi(X_t)] \\ &= \mathbb{E}_t [M_t^*(X_{t+1}) \Psi(X_t)], \end{aligned} \quad (7)$$

For the valuation of an asset with payoffs that are functions of the underlying only, the pricing kernel is also a function of these asset payoffs. We summarise the characteristics of the pricing kernel as follows:

- Univariate function that can vary over time.
- Function that reflects time variation in pricing kernel state variables.

Under the assumption that investors have a finite time horizon over which they invest and the equity index level is equal to aggregate wealth, [Ait-Sahalia and Lo \(2000\)](#) and [Jackwerth \(2000\)](#) estimate pricing kernels projected onto equity return states using equity index option prices. Importantly the authors note that the projected pricing kernel can be interpreted in the same way as the original (although the two are not the same.)

One would interpret the projected pricing kernel as follows:

$M_t^*(X_{t+1})$	
constant	investors are indifferent to a unit payoff across different states
increasing	investors are showing an increasing desire for a unit payoff across states
decreasing	investors are showing a decreasing desire for a unit payoff across states

Now define a measure of risk aversion for the projected pricing kernel that is consistent with the original Arrow-Pratt measure of relative risk aversion defined in equation 3:

$$\gamma_t^* = -\frac{X_{t+1}M_t'^*(X_{t+1})}{M_t^*(X_{t+1})}, \quad (8)$$

Level of projected risk aversion determines the relative preference for a unit payoff across states. High γ_t^* implies that price kernel projections are steeply negatively sloped and that the demand for hedging securities (securities that payoff when asset prices are too low) are high.

3.2 Reconciling pricing methods

The following results are derived (equations 5 and 7) for the pricing of an asset or derivative security using the pricing kernel and projected pricing kernel respectively:

$$p_t = \mathbb{E}_t [M_t(Z_t, Z_{t+1})\Psi(X_t)],$$

$$p_t = \mathbb{E}_t [M_t^*(X_{t+1})\Psi(X_t)],$$

where

$$\begin{aligned} M_t^*(X_{t+1}) &= \int M_t(Z_t, Z_{t+1})f_t(Z_{t+1}|X_{t+1}) dZ_{t+1} \\ &= \mathbb{E}_t [M_t(Z_t, Z_{t+1})|X_{t+1}], \end{aligned}$$

The pricing with $M_t^*(X_{t+1})$ and $M_t(Z_t, Z_{t+1})$ is equivalent if the projection is unique (one to one function). If we assume a linear functional form, $\Psi(X_t) = X_t$, then the projection is unique.

Now consider the risk neutral pricing equation (Bjork, 2004; Shreve, 2004):

$$\begin{aligned}
 p_t &= e^{-rT} \mathbb{E}^Q [X_{t+1}] \\
 &= e^{-rT} \int X_{t+1} g_t(X_{t+1}) dX_{t+1} \\
 &= e^{-rT} \int X_{t+1} \frac{g_t(X_{t+1})}{f_t(X_{t+1})} f_t(X_{t+1}) dX_{t+1}, \tag{9}
 \end{aligned}$$

where:

$$\begin{array}{l|l}
 g_t(X_{t+1}) & \text{Risk neutral density (SPD)} \\
 f_t(X_{t+1}) & \text{Historic density}
 \end{array}$$

Now since equations 6 and 9 are equivalent for any $\Psi(X_t)$, the pricing kernel is easily reconciled to yield:

$$M_t^*(X_{t+1}) = e^{-rT} \frac{g_t(X_{t+1})}{f_t(X_{t+1})}, \tag{10}$$

3.3 Consumption based approach to asset pricing

Consider a two period economy where the investors' utility function over period t and $t + 1$ consumption is defined as follows (Cochrane, 2002):

$$U(C_t, C_{t+1}) = U(C_t) + \beta \mathbb{E}_t [U(C_{t+1})], \tag{11}$$

where:

$$\beta \mid \text{Subjective discount factor}$$

Let ζ be the amount that an investor purchases of risky assets at time t . The investors problem amounts to choosing ζ to maximise $U(C_t, C_{t+1})$:

$$\max_{\zeta} U(C_t) + \beta \mathbb{E}_t [U(C_{t+1})], \tag{12}$$

$$\text{s.t.} \quad \begin{array}{l}
 C_t = e_t - p_t \zeta \\
 C_{t+1} = e_{t+1} + \Psi(X_{t+1}) \zeta,
 \end{array}$$

where:

$$\begin{array}{l|l}
 C_t & \text{Investors consumption in period } t \\
 e_t & \text{Investors endowment in period } t
 \end{array}$$

The projected pricing kernel is equivalent to the original pricing kernel for assets that have payoff functions that depend linearly on X_{t+1} . Under this assumption ($\Psi(X_t) = X_{t+1}$) the condition for optimal consumption and portfolio choice is easily solved to yield the following first order condition:

$$p_t U'(C_t) = \beta \mathbb{E} [U'(C_{t+1}) X_{t+1}],$$

Rearranging terms, find the consumption-based asset pricing model:

$$p_t = \mathbb{E}_t \left[\beta \frac{U'(C_{t+1})}{U'(C_t)} X_{t+1} \right], \quad (13)$$

Then, define the pricing kernel as follows:

$$M_{t+1} = \left[\beta \frac{U'(C_{t+1})}{U'(C_t)} \right], \quad (14)$$

Having defined the pricing kernel this way, it is easily seen that the kernel is a ratio of marginal utilities. As [Cochrane \(2002\)](#) notes, the fundamental measure of how you feel is marginal utility. Assets that payoff well during good times when consumption is high and the marginal utility of consumption is low are less attractive in poorer times when consumption is low and the marginal utility of consumption is high ([Constantinides et al., 2003](#)).

The pricing kernel can now be interpreted as the marginal rate of substitution, namely as the rate at which the investor is willing to substitute consumption at time $t+1$ for consumption at time t . Having an understanding of how an investor interprets the pricing kernel, the following relation is derived that reconciles the density functions with marginal utilities:

$$M_t^*(X_{t+1}) = e^{-r(t+1)} \frac{g_t(X_{t+1})}{f_t(X_{t+1})} = \beta \frac{U'(C_{t+1})}{U'(C_t)}, \quad (15)$$

Having this relationship one can interpret γ_t^* , as defined in equation 8, as a function of marginal utilities.

4 Estimation strategy

4.1 Specification of the state price density

Empirical studies document the result that the prices of index options differ from the Black-Scholes formula price¹. This result, apparent in equity option prices, is due to the volatility smile. The smile observed in the equity options market describes the relationship between the strike price and implied volatility. Prior to the market crash of 1987, implied volatilities were independent of strike prices. However, since 1987, traders have used the volatility smile to price options. As the strike increases, volatility decreases.

Two schools of thought have explained the presence of the smile. [Hull \(2006\)](#) proposes that as a company's equity decreases in value, its leverage (liabilities) increases, implying that equity becomes riskier and hence volatility increases and vice versa. The argument proposed explains that volatility is decreasing function of price. [Rubinstein \(1994\)](#) however proposes the empirical argument that suggests that the likelihood of an extreme event occurring, namely a market crash or loss, is higher than the standard lognormal assumption under Black Scholes. After the market crash of 1987, traders are consequently fearful of another such shock, and hence price options at higher volatilities (prices) with lower strikes and lower maturities. This emphasises the distinction between the historical distribution and the risk neutral distribution of returns, because we are able to note that investors are more fearful about losses than any upside return of the same magnitude.

As we aim to recover the risk neutral distribution from option price data, we need to have a model of the implied volatility smile. Various authors have attempted to capture the effects of this smile by building models that include conditional heteroscedascity and later a leverage parameter². The leverage effect generates a negative skew in stock returns, which together with a stochastic volatility model captures the effect that volatility increases more when the stock price drops. Consequently, authors have found that these models significantly improve upon the performance of the Black Scholes model³

[Christoffersen et al. \(2003\)](#) note that despite the appeal of these models, for options with shorter dated maturities, the models still fail to explain all the biases in option prices. The error distribution in such models is assumed to be normal. However, to account for the volatility smile effect the innovation distribution needs to be characterised with higher non-zero moments. The option valuation model proposed by [Christoffersen et al. \(2003\)](#) yields option prices

¹See references: [Rubinstein \(1985\)](#), [Rubinstein \(1994\)](#), [Jackwerth and Rubinstein \(1996\)](#), [Ait-Sahalia and Lo \(1998\)](#), [Bakshi et al. \(1997\)](#), [Jackwerth \(2000\)](#)

²See references: [Heston \(1993\)](#), [Nandi \(1998\)](#), [Heston and Nandi \(2000\)](#)

³See references: [Bakshi et al. \(1997\)](#), [Ding and Granger \(1996\)](#), [Engle and Mustafa \(1992\)](#), [Jones \(2003\)](#), [Pan \(2002\)](#)

consistent with a return process that contains conditional skewness, conditional heteroscedasticity and a leverage effect. The discrete time Inverted Gaussian GARCH models the conditional return innovation using an Inverse Gaussian distribution and captures the skewness present in short-term as well as long-term returns. When compared to various benchmarks, the model achieves a better fit than standard models in sample and up to 10 weeks out-of-sample. Interestingly, one of the continuous - time limits is the standard stochastic volatility model of [Heston \(1993\)](#).

4.1.1 Stochastic volatility models

Merits of the [Heston \(1993\)](#) model include the ability to model non-lognormal distributions of asset returns as well as leverage effects. The mean reversion property of volatility is captured while the model remains analytically tractable. The implied volatility smiles and surfaces that the model generates are relatively stable and unchanging over time ([Mikhailov and Nogel, 2003](#)).

A caveat when implementing the Heston stochastic volatility model is that while the model may produce results close to empirical observations, it is not possible to construct a riskless portfolio (consisting of the underlying and the hedge) if the volatility of the asset varies stochastically. The reason is that volatility is not a tradeable security in the market.

The [Heston \(1993\)](#) model of stochastic volatility follows:

$$\frac{dX_t}{X_t} = \mu dt + \sqrt{V_t} dW_t^s, \tag{16}$$

$$dV_t = \kappa(\theta - V_t) dt + \sigma\sqrt{V_t} dW_t^q, \tag{17}$$

where:

X_t	Underlying asset price
V_t	Variance
W_t^s, W_t^q	Standard Brownian Motions
κ	Mean reversion parameter
θ	Long run variance
σ	Volatility of variance
ρ	Correlation parameter

Brownian motion processes W_t^s and W_t^q are correlated with parameter ρ such that $dW^s \cdot dW^q = \rho dt$.

Despite the attractiveness of the Heston model, there are a few shortcomings. The Fourier integrals required in the calculation of option prices do not always converge to stable solutions. The model would provide increasing explanatory power across a variety of maturities, by allowing the parameter space to be time

dependent. In addition, the model is unable to capture the skew observed in the market for short dated maturities (Mikhailov and Nogel, 2003).

Realistic models of stock prices need to include jumps or a jump diffusion process. Jumps in stock prices do exist in the market as observed in October 1987, 1997 and most recently in October 2008. In addressing some of the shortcomings of the Heston model, Sepp (2003) includes a jump diffusion process with a stochastic volatility model to model a stock price process. A jump diffusion is beneficial in that the process can explain the volatility smile in a stationary way. The inclusion of a jump diffusion also helps to explain smiles at shorter maturity.

The empirical results in Sepp (2003) support the inclusion of a jump diffusion process. Jump diffusions that are used to model stock price movements reduce squared errors by 72%, while stochastic volatility models reduce errors by 96%. A combination of the two is, hence, expected to yield even better results. We are now able to handle shorter maturities better and improve the overall quality of calibration. Having motivated the inclusion of a jump diffusion process, we include the Heston model with jumps or the model of Bates (1996). The inclusion of a lognormal jump process makes the model tend towards a more realistic match to the data.

The Bates (1996) model of stochastic volatility follows:

$$dX_t = (\mu - \lambda m)X_t dt + \sqrt{V_t}X_t dW^s + (e^J - 1)X_t dN_t, \quad (18)$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW^q, \quad (19)$$

where:

$N(t)$	A Poisson process with constant intensity λ
J	Jump magnitude (random variable) with density function $\bar{w}(J)$
m	Average jump amplitude (defined below)

The risk neutral dynamics proposed above ensure that the discounted stock price process remains a martingale (see Appendix B.1 for proof).

Lognormal jump diffusions:

Merton (1976) proposed a density function for the magnitude of the jump where the logarithm of the jump size is normally distributed with mean v and standard deviation δ .

$$\bar{w}(J) = \frac{1}{\sqrt{2\pi\delta^2}} e^{-\frac{(J-v)^2}{2\delta^2}}, \quad (20)$$

The variable e^J is hence lognormally distributed and $m = e^{v + \frac{1}{2\delta^2}} - 1$

Having identified two models to model the volatility surface, from which we

shall then construct option prices, we proceed to finding a pricing formula using the method of characteristic functions and Fourier transforms. The formula derived below is consistent with the style of the Black Scholes formula. This is a generalised solution following the methodology of [Sepp \(2003\)](#) assuming that the characteristic functions follow an affine term structure (the details are contained in [Appendix C](#)). We confine the analysis to European vanilla call options.

The price of a European call option c :

$$c = XP_1 - Ke^{-(r-q)(T-t)}P_2, \quad (21)$$

P_1 is interpreted as the delta of option, while P_2 is the conditional risk neutral probability that the asset prices ends in the money at maturity ([Mikhailov and Nogel, 2003](#)).

If we have characteristic functions ϕ_1 and ϕ_2 , we can find P_1 and P_2 using inverse Fourier transforms:

$$P_j = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathbb{R} \left[\frac{\phi_j(u)e^{-iulnK}}{iu} \right], du, j = 1, 2 \quad (22)$$

Assuming that the characteristic functions follow an affine structure, we have a general form of the function as follows ([Sepp, 2003](#)):

$$\phi_j(u) = e^{iuS + A(u,\tau) + B(u,\tau)V + C(u,\tau) + D(u,\tau)\lambda}, \quad (23)$$

where:

$$S = \ln \left(\frac{X}{K} \right) + (r - q)\tau,$$

In finding a solution to this pricing problem, we define variables k , I , and b as follows:

	k	I	b
$j = 1$	+1	1	$\kappa - \rho\sigma$
$j = 2$	-1	0	κ

and then proceed to the following solution ([Sepp, 2003](#); [Mikhailov and Nogel, 2003](#)):

Heston Model

1. Stochastic Volatility:

$$A(u, \tau) = -\frac{\kappa\theta}{\sigma^2} \left[\psi_+\tau + 2 \ln \left(\frac{\psi_- + \psi_+ e^{-\zeta\tau}}{2\zeta} \right) \right],$$

$$B(u, \tau) = -(u^2 - kiu) \frac{1 - e^{-\zeta\tau}}{\psi_- + \psi_+ e^{-\zeta\tau}},$$

$$\psi_{\pm} = \mp(b - \rho\sigma iu) + \zeta,$$

$$\zeta = \sqrt{(\kappa - \rho\sigma iu)^2 + \sigma^2(u^2 - kiu)},$$

Bates Model

1. Stochastic Volatility:

$$A(u, \tau) = -\frac{\kappa\theta}{\sigma^2} \left[\psi_+ \tau + 2 \ln \left(\frac{\psi_- + \psi_+ e^{-\zeta\tau}}{2\zeta} \right) \right],$$

$$B(u, \tau) = -(u^2 - kiu) \frac{1 - e^{-\zeta\tau}}{\psi_- + \psi_+ e^{-\zeta\tau}},$$

$$\psi_{\pm} = \mp(b - \rho\sigma iu) + \zeta,$$

$$\zeta = \sqrt{(\kappa - \rho\sigma iu)^2 + \sigma^2(k^2 - kiu)},$$

2. Constant Jump Rate Intensity

$$C(u, \tau) \equiv 0, \quad D(u, \tau) = \tau\Lambda(u).$$

3. Log-Normal jump size distribution

$$\Lambda(u) = e^{(v+I\frac{\delta^2}{2})iu - \frac{\delta^2 u^2}{2} + I(v+\frac{\delta^2}{2})} - 1 - (iu + I)(e^{v+\frac{\delta^2}{2}} - 1),$$

4.1.2 Model calibration

The next step is to calibrate these models to real data. Calibration faces three challenges, namely:

1. Given that we have allowed the volatility to vary stochastically, an exact likelihood function cannot be computed. This means that standard econometric tools that would lead to a best, linear, unbiased estimate of the underlying return distribution are fruitless.
2. The model is posed in a continuous time setting, whereas the data is observed at discrete time intervals.
3. Finally, as noted above, the market is incomplete. Volatility is not a traded instrument and the absence of an observable market price of volatility means that we do not know the volatility risk premium when pricing options.

Despite these challenges, we proceed through an implied estimation strategy. This strategy may lack the foundation of a fundamental statistical theory; however, we proceed in a similar fashion to Bakshi et al. (1997), Nandi (1998) and Lin et al. (2001) who estimate implied parameters from cross sectional option data. The solution is to find parameters for the models, which produce “correct” market prices of vanilla options. To do this we use the models’ fit into the implied volatility surface as a measure of goodness of fit. Sepp (2003) suggests the following:

$$\min_{\theta} \sum_{j=1}^N \left(\sigma_j^{\S} - \sigma_j^{\dagger}(\theta) \right)^2, \quad (24)$$

where:

$$\begin{array}{l|l} \theta & \text{Parameter space} \\ \sigma_j^{\S} & \text{Market implied volatility} \\ \sigma_j^{\dagger} & \text{Model volatility} \end{array}$$

Here we minimise the squared differences between the market volatility and the model volatility over the parameter space θ .

Another measure of goodness of fit is one that minimises the squared differences between prices:

$$\min_{\theta} \sum_{j=1}^N \left(c_j(\sigma_j^{\S}) - c_j(\sigma_j^{\dagger}; \theta) \right)^2, \quad (25)$$

where:

$$\begin{array}{l|l} c_j(\sigma_j^{\S}) & \text{Market price of call options} \\ c_j(\sigma_j^{\dagger}) & \text{Model price of call options} \end{array}$$

Now equations 24 and 25 are equivalent if we use the same model to price the call options. Herein we can introduce our models as defined above. If the model parameters produce a volatility surface close to that of the observed market surface, we will produce option prices close to the market price of options.

Equation 25 is a least squared error problem; however as Becker (2009) and Moodley (2005) indicate, this is also an inverse problem as we solve for parameters indirectly through an implied structure (implied estimation). Inverse problems are generally not well posed and there may be very sensitive to parameter estimates. These solutions are also subjective to the initial guess.

Finding a global minimum is difficult and time consuming. It is dependent on the actual calculation used in the optimisation procedure. Hence, any solution to an inverse problem may be questionable. We have an ill-posed, inverse problem that needs to be solved (Moodley, 2005).

Our aim would be to select a ‘reasonable’ initial guess and add some element of regularisation to ensure that the system achieves stability. Introduce a penalty function $p(\theta)$ to our problem to regularise the problem (by ensuring that the objective function is convex) (Chiarella et al., 2000):

$$\min_{\theta} \sum_{j=1}^N \left(c_j(\sigma_j^{\S}) - c_j(\sigma_j^{\dagger}; \theta) \right)^2 + \alpha p(\theta), \quad (26)$$

We now have a problem that we can solve that will yield a solution close to the true solution, but is at the same time practically implementable. The penalty we introduce follows a hybrid version of that suggested by Mikhailov and Nogel (2003):

$$\alpha p(\theta) = \|\theta - \theta_0\|^2, \quad (27)$$

where:

$$\theta_0 \mid \text{Initial parameter estimate}$$

This penalty function introduces stability to the system. However, we also use the mean square error (MSE) as a measure goodness of fit.

If the MSE is minimised and the penalty function is satisfied, we proceed with our next iteration of the parameter estimates to minimise the MSE again. If the penalty function is violated, the procedure stops and reverts to the previous parameter estimate. This problem is solved in MATLAB using the `lsqnonlin` function (see appendix D for details on the function).

4.1.3 Simulating the risk neutral distribution

Having understood the nature of the stochastic volatility models that we propose to fit to the data and having calibrated these models to the data, we are now well positioned to make a reasonable assumption about the underlying price process suggested by the model as well as the nature of the underlying distribution constructed from these models. For purposes of this paper, these distributions will be assumed the best approximations of the risk neutral distributions.

The thought process is summarised as follows:

1. Market prices of liquidly traded option prices are difficult/expensive to obtain.
2. Use the next best alternative of available information, namely the implied volatility surface.
3. Fit two models (Heston/Bates) to the data and find an acceptable level of goodness of fit so that we can be satisfied that the models approach reality.

4. Construct ‘market’ prices from these models.
5. From these market prices obtain the risk neutral distribution.
6. Use this in the calculation of the empirical pricing kernel.

Steps 4 and 5 are eliminated as we can calculate the distribution of the stock price as a Monte Carlo exercise (see appendix E for details).

4.2 Specification of the historical density

4.2.1 Lognormal density function

Black and Scholes (1973) developed the most common methodology for valuing the price of European style derivatives. Their model assumes that under the martingale measure (risk-neutral measure) \mathbb{Q} , an asset price follows a Geometric Brownian Motion with constant drift (μ) and constant diffusion (σ).

$$\frac{dX_t}{X_t} = \mu dt + \sigma dW_t, \quad (28)$$

Under these assumptions, we can easily derive that the stock price process is lognormally distributed and that log-share prices follow a normal distribution with mean $(\mu - \frac{\sigma^2}{2})(T - t)$ and variance $\sigma^2(T - t)$. To estimate the density function of stock returns, we easily derive:

$$\ln\left(\frac{X_T}{X_t}\right) = r_X \sim N\left(\left(\mu - \frac{\sigma^2}{2}\right)(T - t), \sigma^2(T - t)\right), \quad (29)$$

where μ and σ are estimated from historic data.

We proceed by sampling from a normal distribution and then calculating the relevant stock prices to sample from a lognormal distribution.

4.2.2 Empirical innovation density function

The second model that we propose is a stochastic volatility model that incorporates salient features of the equity index return process. We use a *ARMA*(1, 1) model to model the stock return process and a *EGARCH*(1, 1) model for the volatility of the error distribution⁴. The choice behind the EGARCH model is to capture the leverage effect.

We have a sequence of stock returns:

$$r_{X_t} = \ln\left(\frac{X_T}{X_t}\right), \quad (30)$$

⁴We use the GARCH and EGARCH terms interchangeably in this paper as EGARCH models form a subset of GARCH models.

Now we assume that the process can be modeled as follows:

$$r_{X_t} = \eta + \alpha_1 r_{X_{t-1}} + \beta_0 \epsilon_t + \beta_1 \epsilon_{t-1}, \quad (31)$$

where we assume ϵ_t is drawn from an **empirical** distribution function with stochastic variance:

$$\epsilon_t \sim f(0, \sigma_{t|t-1}^2),$$

and we have the following EGARCH process:

$$\log(\sigma_{t|t-1}^2) = \gamma + \alpha_0 g(\epsilon_{t-1}^2) + \beta_0 \log(\sigma_{t-1|t-2}^2), \quad (32)$$

where $g(\epsilon_t) = \theta_0 \epsilon_t + \gamma_0 (|\epsilon_t| - \mathbb{E}[|\epsilon_t|])$.

To model the density function of the disturbances f , we use a filtered historic sampling methodology. This non-parametric methodology extracts the residuals and conditional variances from the EGARCH model (using historic data) and constructs the series of independent and identically distributed standardised residuals. The nature of the underlying process, as captured by the residuals are then used to construct a bootstrapped series of residuals from which one can simulate the asset path progression.

4.2.3 Nonparametric density function

In liquid markets, one is able to estimate the volatility skew or surface from traded option prices. The risk neutral distribution can also be estimated from the price of derivative securities; it is an implicit assumption that there is sufficient information available for us to extract this distribution. This paper uses stochastic volatility models to construct ‘market’ prices of options and then determine the risk neutral distribution.

[Stutzer \(1996\)](#) and [Duan \(2002\)](#), however provide an alternative approach that utilises non-parametric techniques and numerical methods for option pricing and estimation of the volatility skew. [Duan \(2002\)](#) suggests a methodology using the relative entropy principle and information on the underlying asset only.

Non-parametric approaches impose no prior assumptions about the nature of the underlying distribution of the asset. Simplifying assumptions in a parametric setting can often detract from the true results and hence the motivation for the inclusion of this non-parametric model. For proposes of this paper, however, we shall confine a discussion of estimation of this density function as the historical density function.

Non-parametric option pricing techniques use spot market data to determine the probability distribution of the underlying asset. [Derman and Kani \(1994\)](#) and [Rubinstein \(1994\)](#) match estimated prices to market prices using implied binomial trees. We, however, follow the canonical valuation methods of [Duan \(2002\)](#) and [Stutzer \(1996\)](#). We are able to gain information about the volatility

skew, option prices and hence the underlying distribution of asset returns in markets where there is a lack of information on traded option prices or a lack of liquidity.

We follow the methodology outlined in [Duan \(2002\)](#) and [Arajo and Mar \(2006\)](#) to find the density function. Given asset returns r_i with a sample mean $\bar{\mu}$ and standard deviation $\bar{\sigma}$, we define the normalised asset return z_i with associated distribution function $g(\cdot)$ as follows:

$$z_i = \Phi^{-1} \left(g \left(\frac{r_i - \bar{\mu}}{\bar{\sigma}} \right) \right), \quad (33)$$

where:

Φ | Standard normal distribution function

To estimate $g(\cdot)$ we proceed by estimating the empirical distribution function of the standardised return $\left(\frac{r_i - \bar{\mu}}{\bar{\sigma}} \right)$ namely:

$$g(\hat{x}) = \frac{1}{N} \sum_{i=1}^N \left(\frac{r_i - \bar{\mu}}{\bar{\sigma}} \leq x \right), \quad (34)$$

We use kernel smoothing to smooth the function to find $g(\cdot)$.

To find our (risk - neutral) distribution function we need to find a density function $f(x)$ that minimises the relative entropy between itself and $\phi(x)$ (the standard normal density) subject to the constraint that the expected return of the distribution is equal to the risk free rate less the dividend yield.

Using the relative entropy principle, the density function for z_i is given by minimising the following integral:

$$\int_{-\infty}^{\infty} f(x) \ln \left(\frac{f(x)}{\phi(x)} \right) dx, \quad (35)$$

subject to the constraints:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) dx &= 1, \\ \int_{-\infty}^{\infty} x f(x) dx &= a_t, \end{aligned}$$

where:

a_t | Constant that is a function of time

a_t allows us the freedom to let the density function be a function of time. This problem has solution [Cover and Thomas \(1991\)](#) (see Appendix [B.2](#) for proof):

$$f(x, \lambda_t) = \frac{\phi(x)e^{\lambda_t x}}{\int_{-\infty}^{\infty} \phi(x)e^{\lambda_t x} dx} = \phi(x - \lambda_t), \quad (36)$$

The value of λ_t corresponds to the value of a_t and hence we have a density function that is fully parameterised by λ_t . λ_t can be solved using numerical methods, subject to the constraints defined above. The risk neutral density of the normalised returns is now given by $\phi(x - \lambda_t^*)$, from which we derive the stock price process:

$$X_t = X_{t-1} e^{\sigma_t g^{-1}(\Phi(z_t)) + \mu_t}, \quad (37)$$

where:

$$z_t \mid \text{Normal random variable with mean } \lambda_t \text{ and variance } 1$$

The asset price until maturity may be modeled as a Monte Carlo exercise.

4.3 Estimating the pricing kernel

The accuracy of the pricing kernel can be judged by the extent to which it reproduces the price of traded assets. We select the empirical pricing kernel as the function that provides the best fit to derivative prices, given current expectations about future profits or payoffs. In our case our derivative prices are the prices calculated from our stochastic volatility models. Our empirical pricing kernel is an estimate of the pricing kernel projection, $M_t^*(X_{t+1})$, on a particular date, rather than as an average estimate over time.

Recall the following from section [3.2](#):

$$p_t = \mathbb{E}_t [M_t^*(X_{t+1})\Psi(X_t)],$$

where

$$M_t^*(X_{t+1}) = \mathbb{E}_t [M_t(Z_t, Z_{t+1}) | X_{t+1}],$$

Let us rewrite the pricing equation for a derivative with a payoff that depends on the return of the underlying asset(r_{t+1}), ([Rosenberg and Engle, 2002](#)):

$$\begin{aligned} p_{i,t} &= \mathbb{E}_t [M_t^*(r_{t+1})g_i(r_{t+1})] \\ &= \int M_t^*(r_{t+1})g_i(r_{t+1})f_t(r_{t+1})dr_{t+1}, \end{aligned} \quad (38)$$

where:

$$\begin{array}{l|l} p_{i,t} & \text{Price of } i^{th} \text{ asset (derivative)} \\ g_i & \text{Derivative payoff function} \\ f_t & \text{Probability density of one period underlying asset returns} \end{array}$$

In addition, we note that $M_t^*(r_{t+1})$ is an implicit function of prices ($p_{i,t}$), payoffs ($g_i(r_{t+1})$) and probabilities ($f_t(r_{t+1})$).

We now rewrite the pricing equation to find a formula for fitted asset prices $\widehat{p}_{i,t}$ using an estimated pricing kernel projection $\widehat{M}_t^*(r_{t+1})$ and estimated payoff density function $\widehat{f}_t(r_{t+1})$:

$$\begin{aligned}\widehat{p}_{i,t} &= \mathbb{E}_t \left[\widehat{M}_t^*(r_{t+1}) g_i(r_{t+1}) \right] \\ &= \int \widehat{M}_t^*(r_{t+1}) g_i(r_{t+1}) \widehat{f}_t(r_{t+1}) dr_{t+1},\end{aligned}\quad (39)$$

Now estimate the pricing kernel projection as a best fit function, namely as the function that makes model (fitted) prices closest to market prices using the estimated payoff density. This problem is easily simplified by letting the kernel projection be a parametric function. We follow the definition of [Rosenberg and Engle \(2002\)](#) for the functional form of the pricing kernel:

$$M_t^*(r_{t+1}) = \theta_{0,t}(r_{t+1})^{-\theta_{1,t}}, \quad (40)$$

The significant advantage of assuming this power form is that we can interpret $\theta_{1,t}$ as a measure of empirical risk aversion over time (see appendix [A.1](#) for derivation).

To find this projected pricing kernel we need to solve the following problem:

$$\min_{\theta_t} \sum_{i=1}^N [p_{i,t} - \widehat{p}_{i,t}(\theta_t)]^2, \quad (41)$$

We estimate $\widehat{p}_{i,t}(\theta_t)$ through Monte Carlo simulation. We simulate H returns, and estimate the integral in equation [39](#) as follows:

$$\begin{aligned}\widehat{p}_{i,t}(\theta_t) &= \frac{1}{H} \sum_{h=1}^H [M_t^*(r_{t+1,h}; \theta_t) g_i(r_{t+1,h})] \\ &= \frac{1}{H} \sum_{h=1}^H [\theta_{0,t}(r_{t+1,h})^{-\theta_{1,t}} g_i(r_{t+1,h})],\end{aligned}\quad (42)$$

$\theta_{0,t}$ and $\theta_{1,t}$ are easily solved through an optimisation procedure. We use similar numerical methods as in section [4.1.2](#).

5 Data

5.1 S&P500 equity index

The S&P 500 Index consists of 500 stocks traded on stock exchanges throughout the United States. The index is constructed from large-cap companies and is noted as the most representative index of market performance in the United States. A particularly appealing aspect of this index is that it has a liquid derivative market.

We use daily end of day closing prices on the index from 2 January 2000 to 31 December 2009. We note some characteristics of the data in table 1 and the underlying performance is shown in figure 1:

	Price	Returns
Observations	2 515	2 514
Mean	1 187.26	-0.01%
Standard deviation	198.00	1.40%
Annualised volatility	-	22.24%
Maximum	1 565.15	0.110
Minimum	676.53	-0.095
Range	888.62	0.204

Table 1: Empirical characteristics of the S&P500 Equity Index

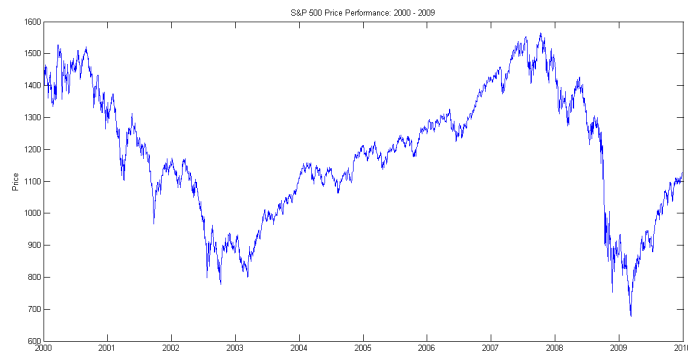


Figure 1: S&P500 Equity Index

5.2 S&P500 implied volatility data

In section 4.1.2 we understand how to fit a stochastic volatility model to an empirical volatility surface. In addition, we understand how to use such a model to price options which we then assume are the market prices of options traded on the underlying index.

The implied volatility data described below is extracted from traded option prices on the S&P500 Index. Given that we are unable to access actual traded option prices, we use this volatility surface as a (realistic) substitute for the market. Table 2 and figure 2 describe the term structure of volatility as function of strike (percentage moneyness) and time (maturity). Our sample period for our study on empirical risk aversion is 2 January 2005 to 31 December 2009. We obtain the volatility surface daily over this period from Bloomberg.

		Maturity					
		30 days	60 days	3M	6M	12M	18M
% Moneyness	80.0%	40.11	38.17	37.38	35.59	33.67	32.72
	90.0%	32.17	32.06	31.98	31.58	30.83	30.37
	95.0%	28.73	29.21	29.51	29.73	29.49	29.25
	97.5%	27.14	27.88	28.33	28.84	28.85	28.70
	100.0%	25.65	26.62	27.22	27.98	28.21	28.17
	102.5%	24.30	25.43	26.15	27.15	27.59	27.64
	105.0%	23.42	24.39	25.13	26.34	26.99	27.13
	110.0%	23.29	22.78	23.24	24.82	25.83	26.13
	120.0%	23.29	20.70	20.38	22.31	23.70	24.30

Table 2: Volatility term structure on 10 June 2009

5.3 S&P500 dividend yield

Dividend yield is measure of the cash return for every dollar you have invested into a stock. The dividend yield on the S&P500 index is calculated by Standard and Poor's by aggregating all dividends paid by companies over a certain period, divided by the price level of the equity index. This methodology uses discrete data and converts it to a continuous series. The reason behind formulating the data in this way is purely for the ease of which it can be substituted into our option pricing formula. Figure 3 shows the dividend yield on a daily basis from 2 January 2005 to 31 December 2009. This data is obtained from Bloomberg.

5.4 Risk free interest rates in the United States

In lieu of our option pricing, we require risk free interest rates or proxies thereof to price options of varying maturities. We use US T-Bill data as a measure of this riskless rate over the period 2 January 2005 to 31 December 2009. Table 3

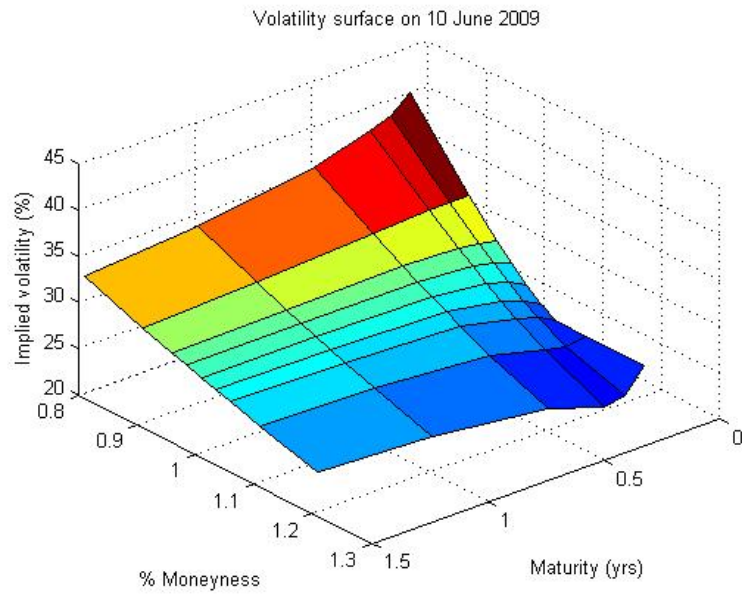


Figure 2: Volatility Surface on 10 June 2009



Figure 3: Dividend yield 2005 - 2009

shares some descriptive statistics while figure 4 provides a graphical display of the risk free rates. Notably, we do not have T-Bill data for the 18 month period, hence we use the T-Bill data for the 12 month maturity as an approximation of the 18 month maturity risk free rate. This data is obtained from Bloomberg.

	30 days	60 days	3M	6M	12M	18M
Mean	2.64%	2.66%	2.72%	2.87%	3.62%	3.62%
Maximum	5.17%	5.07%	5.02%	5.09%	5.76%	5.76%
Minimum	0.00%	0.00%	0.00%	0.09%	0.45%	0.45%
Range	5.17%	5.07%	5.02%	5.00%	5.31%	5.31%

Table 3: Empirical characteristics of the US T-Bill data

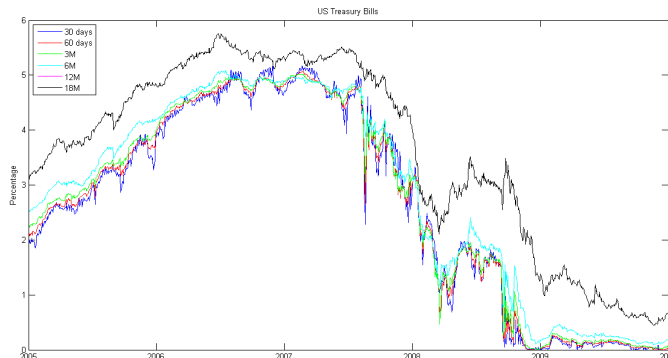


Figure 4: US Treasury Bills Data

5.5 VIX index

The VIX index is a key measure of market expectations of near-term volatility conveyed by the S&P 500 stock index option prices. The VIX is been considered to be the barometer of investor sentiment and market volatility (CBOE, 2010).

This index is a forward looking index constructed from the liquid derivatives (calls and puts) on the S&P500 index. We shall use this data as a measure of “investor fear” in our regression analysis of the empirical risk aversion parameter in section 6.4.1. This data is obtained from Bloomberg.

5.6 TED spread

The TED spread is constructed as the difference between the 3-month T-Bill rate and 3-month LIBOR. It is widely used in the market as measure of inter-

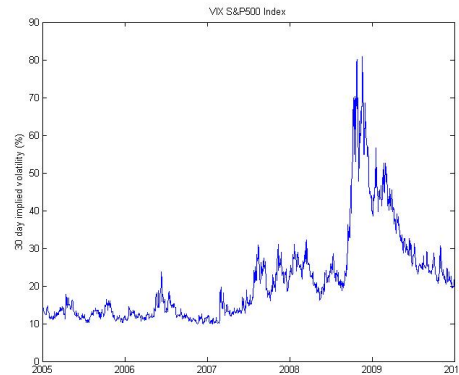


Figure 5: The VIX Index

bank credit risk. We shall use this data as a measure of the price of liquidity in our regression analysis of the empirical risk aversion parameter in section 6.4.1.

The TED spread is measure in basis points and an increase in the TED spread reveals that credit risk or default risk is increasing. In such an instance an investor would typically require a safe investment or hedging security. The TED spread from 2 January 2005 to 31 December 2009 is shown in figure 6. This data is obtained from Bloomberg.

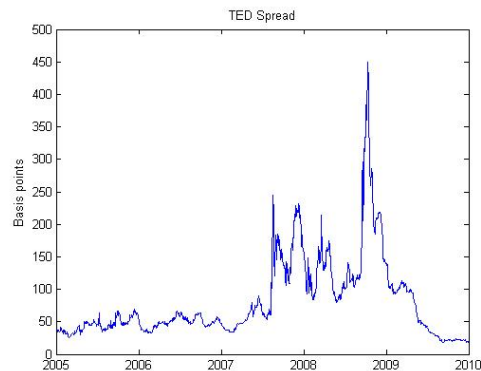


Figure 6: The TED spread

6 Results

6.1 Summary

We are able to fit and calibrate a stochastic volatility model to the market implied volatility surface and then use this model to extract the risk neutral density and “market” option prices. The Bates model provides a satisfactory fit capturing salient features of the market surface. In addition, we are able to find pricing kernel estimates using our density function specifications of historic returns. The pricing kernel exhibits its time varying nature, from which we are able to extract relative risk aversion estimates consistent with investor behaviour and the state of the overall economy. We find the the normal and non-parametric specification for the historic density and the Bates model for the risk neutral density provide the best results.

6.2 Stochastic volatility model calibration

Over the period 2 January 2005 to 31 December 2009, we calibrate the Heston and Bates stochastic volatility models to the market implied volatility surface. We solve for optimal parameters through an implied estimation strategy and hence our initial choice of parameters and bounds for these parameters are an integral element of the optimisation procedure.

Given that we obtain the dimensions of our volatility surface across time and moneyness, we use the percentage moneyness to calculate a “theoretical” strike price each day ($S_t \times \% \text{ moneyness}$). We use this strike price in the calculation of our option prices in the calibration procedure.

Daily we have 54 observations (market-implied volatilities). We aim to minimise the distance between our model volatilities and market volatilities, hence each day we calculate 54 option prices from our model as inputs in our implied estimation procedure.

The benefit of using a generalised methodology to fit a model to a surface and then to extract the risk neutral distribution from simulation of this model, means that each day we have sufficient observations (implied volatilities) to match to our model volatilities. We avoid having to eliminate arbitrage violations from the data as the market implied volatility surface from **Bloomberg** is already constructed with non-negative volatilities.

Furthermore, the time to the maturity of options is determined by the term structure of the surface. Short dated and long dated options that maybe sensitive to liquidity-related biases are eliminated ([Zhang and Shu, 2003](#)). We consider options with maturities ranging from 30 days to 18 months. Options that are out the money or deep in the money are also avoided, as our percentage moneyness is restricted to the 80% - 120% range.

In table 4 we describe our initial guess on 2 January 2005 and the parameter bounds which apply over the entire sample period for the Heston model. From 3 January 2005, the initial guess for the current day’s optimisation is the previous day’s optimal estimate.

		Initial	Lower bound	Upper bound
Kappa	κ	3	2	4
Theta	θ	20%	0%	100%
Sigma	σ	10%	10%	500%
Rho	ρ	-0.5	-1	0
V0	V_0	25%	0%	100%

Table 4: Initial parameter estimate and parameter bounds - Heston model

Our choice of parameters follow the results achieved by [Zhang and Shu \(2003\)](#) when they fit a Heston model to S&P500 data over the period 2 January 1995 to 9 December 1999. Although our sample period is significantly different, the results from [Zhang and Shu \(2003\)](#) are merely used to suggest an initial guess on 2 January 2005. The choice of our parameter bounds are of more interest as they restrict our optimisation procedure to a local optimisation procedure.

The bounds for κ are inspired by [Bergomi \(2005\)](#). θ and V_0 , the long run average variance and initial variance are restricted to between 0.1% and 100%. We know that volatility can never be negative and a historic analysis of S&P500 returns over our sample period suggests that an upper bound of 100% is a reasonable assumption. The volatility of volatility, σ , varies between 10% and 500%. These bounds are not overly restrictive and we maintain the non-negativity of volatility. We restrict the estimate of ρ to between -1 and 0 . The negative correlation parameter aims to capture the “leverage” effect present in equity markets, namely that markets react more to bad news than good news. The negative correlation restriction is consistent with results by [Zhang and Shu \(2003\)](#); [Bakshi et al. \(1997\)](#); [Nandi \(1998\)](#).

In addition to the parameters in the Heston model we describe in table 5 parameters in the Bates model. We let the jump rate intensity in our Poisson process vary between 0 and 25. The mean of the log jump size is restricted between -0.5% and 0.5% which translates to a jump size that is between 60% and 165% of the stock price.

Our parameters estimates in figures 8 and 10 show their stability from 2005 to 2006. From 2007 to 2009, we see that parameter estimates are somewhat more volatile. A plausible explanation could be that the increased volatility in the

		Initial	Lower bound	Upper bound
Jump rate intensity	λ	1	0	25
Log jump size mean	δ	-25%	-50%	50%
Log jump size volatility	v	25%	0%	100%

Table 5: Initial parameter estimate and parameter bounds - Bates model

market meant that the volatility surface showed frequent significant changes. For our models to fit these surfaces, the parameter estimates were not relatively constant. Given our optimisation procedure, however, we are confident that the results yield model surfaces consistent with the market surface over the entire period.

The calibration of the Heston model over our sample period takes approximately 7 days, while the approximation of the Bates model takes approximately 9 days. It is clear that such calibration procedures are computationally intensive. We randomly select the 10 June 2009 to examine our results. It is clear from figures 7 and 9 that the Bates model gives lower percentage differences of the model surface to the market surface compared the Heston model. The Bates model is also able to capture effects of the market skew at lower maturities.

Each day the MSE is calculated. The cumulative MSE over the period is 148.36 and 119.95 for the Heston and Bates models respectively. Given the lower cumulative MSE achieved by the Bates model over our 5-year sample period, we are inclined to accept this model as the dominant result. We include results from our Heston model for sake of completion.

Confident that our calibration procedure ensures optimal parameters over time, we feel that option prices constructed from these stochastic volatility models provide a sufficiently close approximation to options actively traded in the market.

To find our risk neutral distribution we are able to simulate returns from these stochastic volatility models every month. Our daily parameter estimates from our model calibration procedure are averaged each month from which we can simulate returns to find our risk-neutral density function.

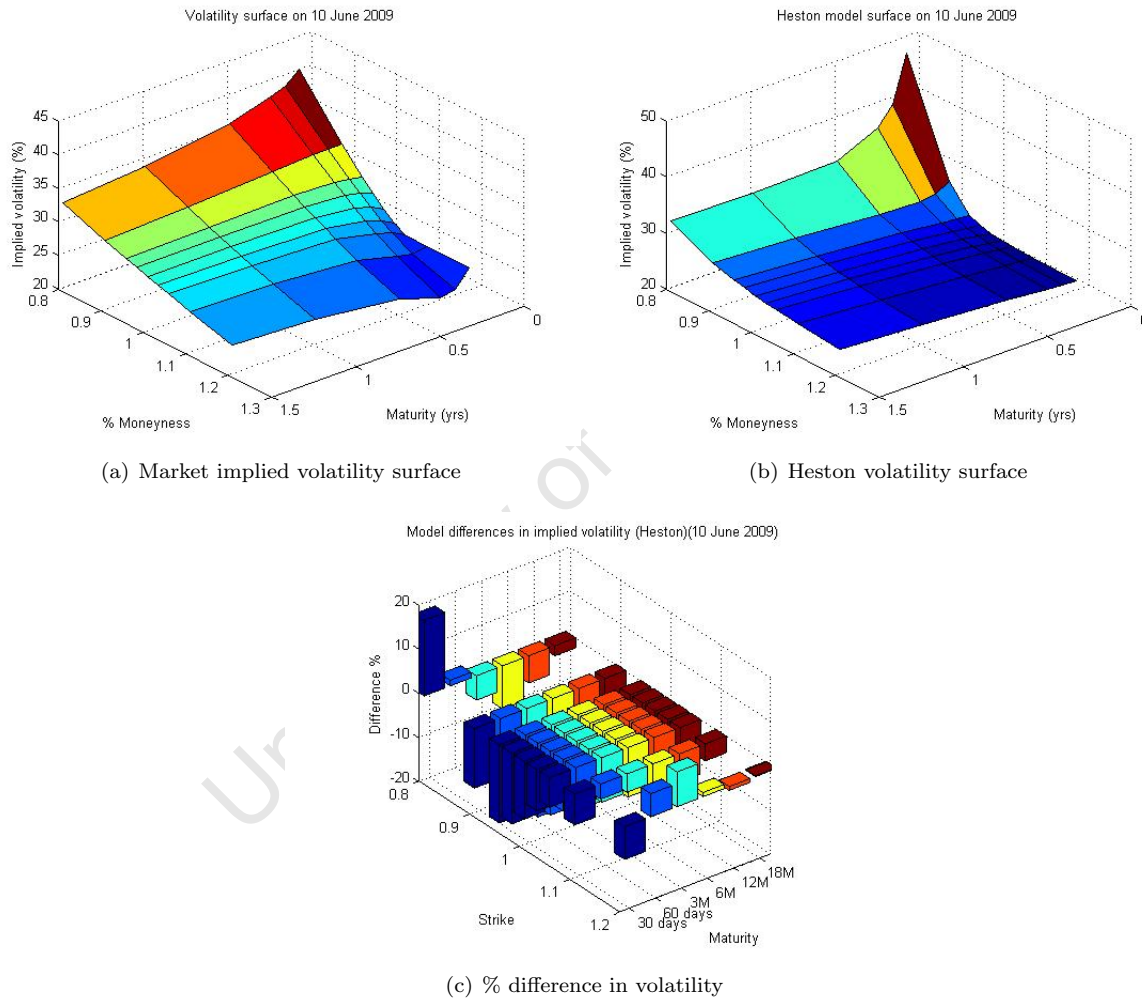
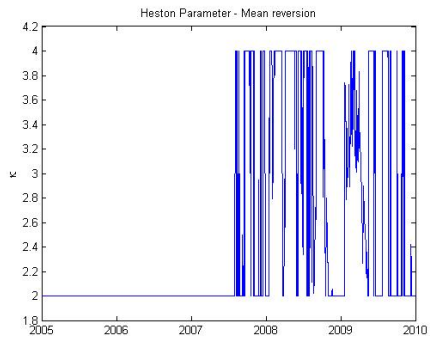
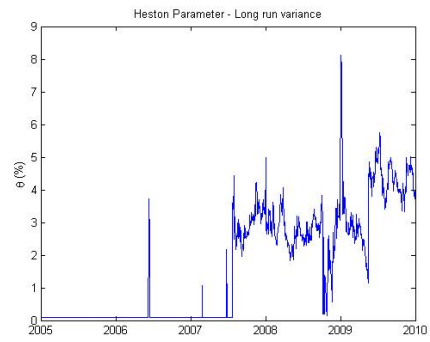


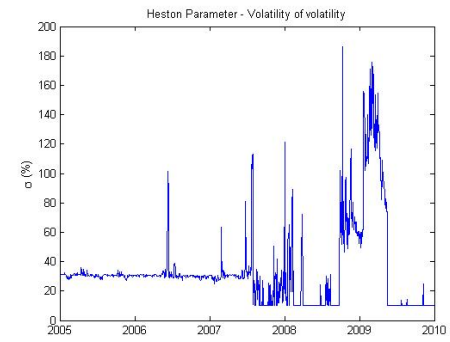
Figure 7: Performance of the Heston Model - 10 June 2009



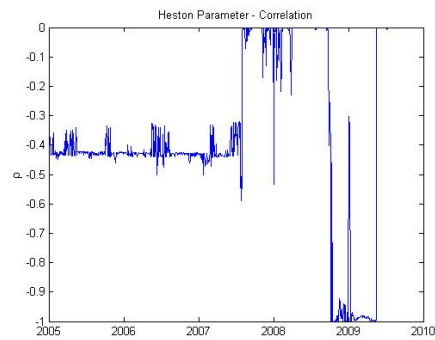
(a) Kappa



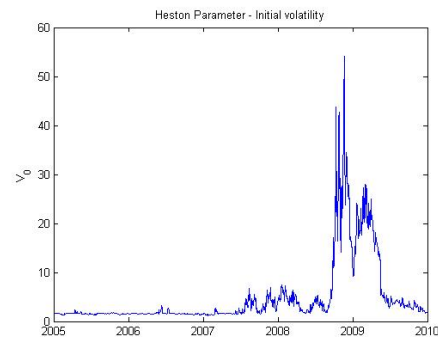
(b) Theta



(c) Sigma

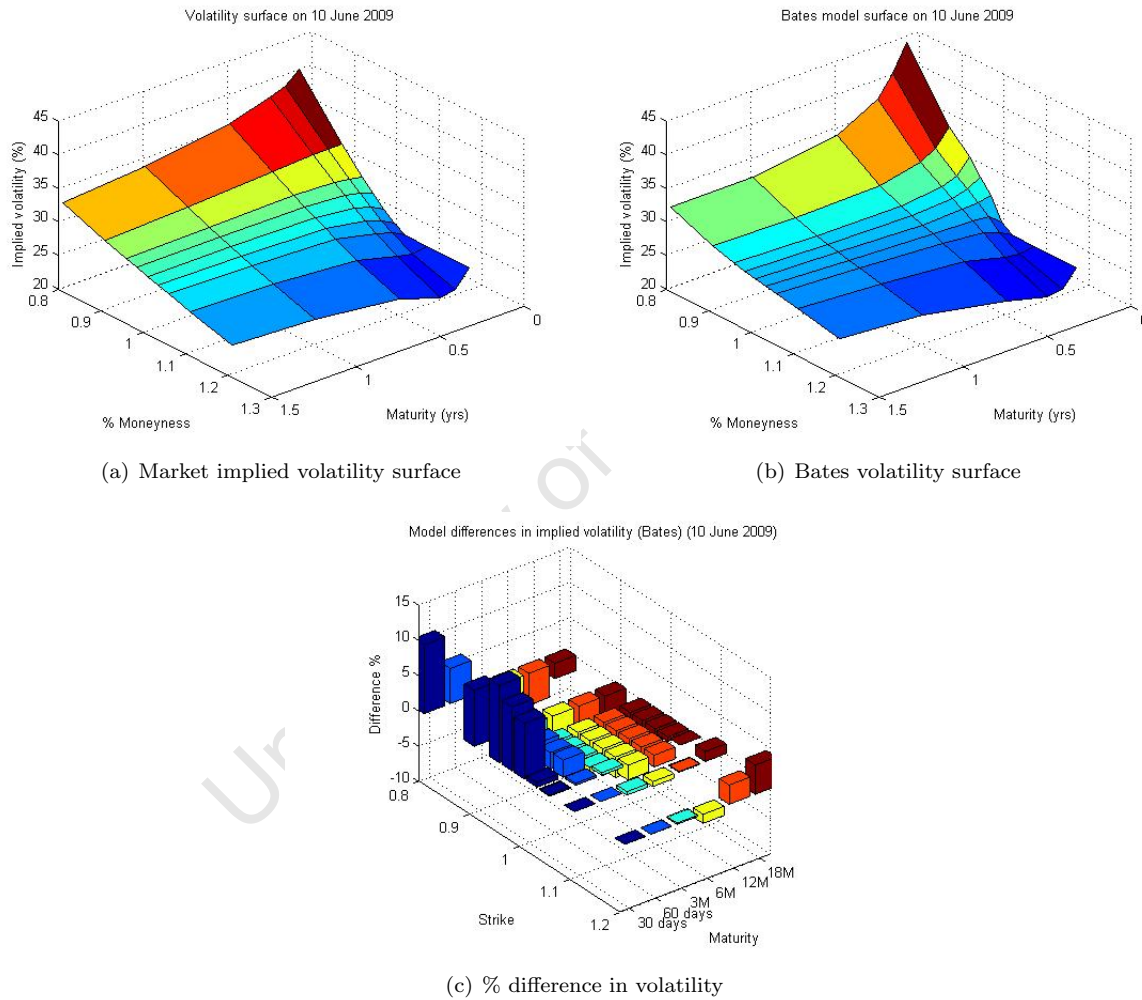


(d) Rho



(e) V0

Figure 8: Parameters in the Heston Model



(a) Market implied volatility surface

(b) Bates volatility surface

(c) % difference in volatility

Figure 9: Performance of the Bates Model - 10 June 2009

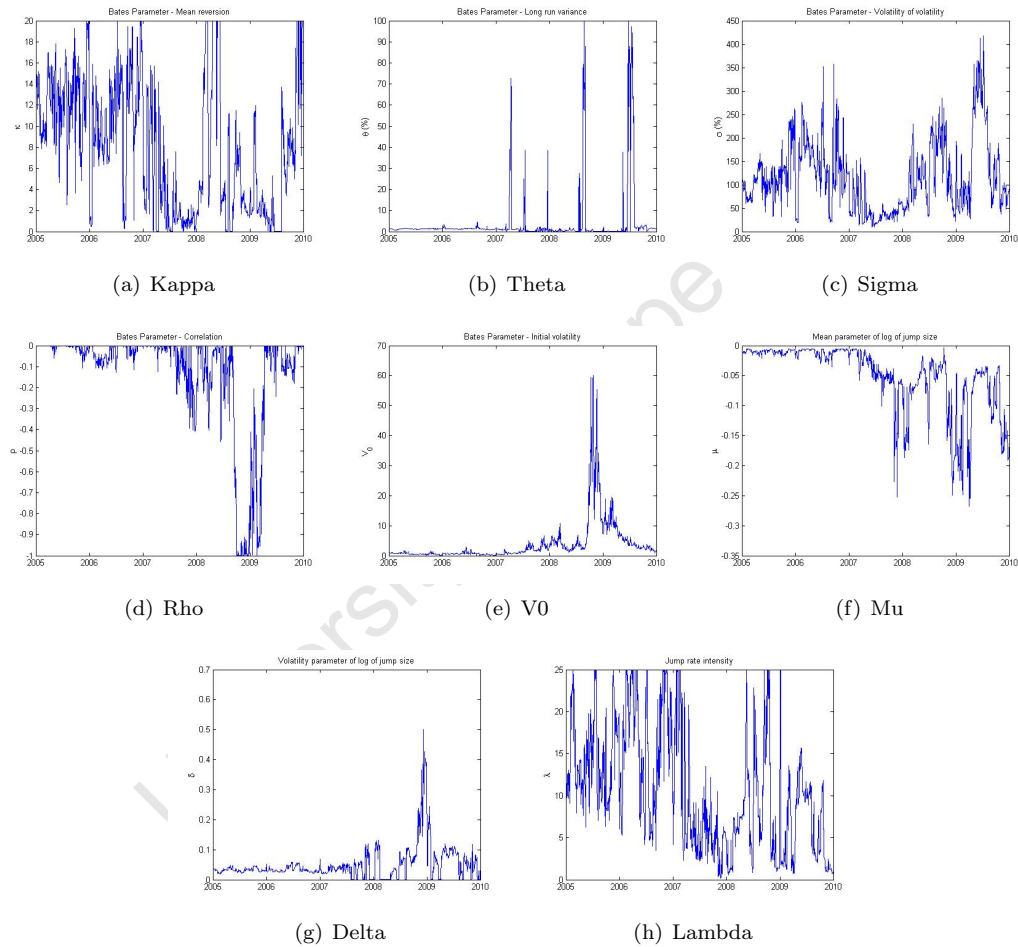


Figure 10: Parameters in the Bates Model

6.3 Historical density functions

6.3.1 Lognormal density function

To calculate parameters in our normal density specification for returns (implying that stock prices evolve according to a lognormal process) we use 5 years of historic return data ($5 \times 252 = 1260$ observations) preceding the day of estimation. We calculate the mean and variance of the historic returns as inputs and specify our density function daily. The parameters are summarised as quarterly averages in table 6.

	Daily mean	Daily Volatility	Annualised Volatility
2005			
Qtr1	-0.01%	1.26%	20.04%
Qtr2	-0.02%	1.22%	19.34%
Qtr3	-0.02%	1.19%	18.89%
Qtr4	-0.01%	1.18%	18.66%
2006			
Qtr1	0.00%	1.14%	18.03%
Qtr2	0.00%	1.09%	17.24%
Qtr3	0.01%	1.07%	16.99%
Qtr4	0.02%	1.03%	16.31%
2007			
Qtr1	0.02%	1.01%	16.02%
Qtr2	0.02%	0.99%	15.75%
Qtr3	0.04%	0.93%	14.81%
Qtr4	0.04%	0.85%	13.47%
2008			
Qtr1	0.04%	0.83%	13.25%
Qtr2	0.03%	0.83%	13.25%
Qtr3	0.02%	0.85%	13.51%
Qtr4	0.00%	1.14%	18.07%
2009			
Qtr1	-0.03%	1.37%	21.82%
Qtr2	-0.02%	1.47%	23.33%
Qtr3	-0.01%	1.50%	23.78%
Qtr4	-0.01%	1.51%	23.98%

Table 6: Estimated parameters in the lognormal density specification

6.3.2 Empirical innovation density function

In our GARCH density specification, we follow the work of [Rosenberg and Engle \(2002\)](#) and [Detlefsen et al. \(2007\)](#) and fit an ARMA model for the return process. We chose to use an EGARCH process to model the volatility of the

empirical distribution. The filtered historic sampling methodology affords us the opportunity to simulate residuals and consequently returns from a bootstrapped matrix of residuals. We use 5 years of historic return data ($5 \times 252 = 1260$ observations) preceding the day of estimation to find our density function of returns. Results of our estimation are summarised in table 7.

	ARMA(1,1)			EGARCH(1,1)			
	Constant	AR(1)	MA(1)	Constant	GARCH(1)	ARCH(1)	Leverage
2005							
Qtr1	0.000	-0.084	0.094	-0.113	0.988	0.108	-0.084
Qtr2	0.000	-0.067	0.076	-0.113	0.988	0.106	-0.084
Qtr3	0.000	-0.349	0.351	-0.114	0.988	0.108	-0.084
Qtr4	0.000	-0.364	0.365	-0.113	0.988	0.107	-0.083
2006							
Qtr1	0.000	-0.119	0.125	-0.113	0.988	0.107	-0.084
Qtr2	0.000	-0.371	0.371	-0.110	0.988	0.106	-0.083
Qtr3	0.000	-0.161	0.166	-0.113	0.988	0.106	-0.085
Qtr4	0.000	-0.190	0.196	-0.114	0.988	0.107	-0.086
2007							
Qtr1	0.000	0.097	-0.099	-0.109	0.989	0.105	-0.083
Qtr2	0.001	-0.053	0.102	-0.735	0.923	0.214	-0.126
Qtr3	0.001	-0.325	0.356	-0.529	0.945	0.178	-0.114
Qtr4	0.001	-0.876	0.871	-0.117	0.988	0.107	-0.089
2008							
Qtr1	0.001	-0.692	0.688	-0.117	0.988	0.107	-0.090
Qtr2	0.001	-0.861	0.855	-0.119	0.988	0.107	-0.091
Qtr3	0.000	-0.333	0.320	-0.115	0.988	0.105	-0.087
Qtr4	0.000	0.714	-0.741	-0.108	0.989	0.111	-0.085
2009							
Qtr1	0.000	0.711	-0.738	-0.108	0.989	0.113	-0.086
Qtr2	0.000	0.708	-0.736	-0.099	0.990	0.111	-0.084
Qtr3	0.000	0.707	-0.735	-0.096	0.990	0.109	-0.084
Qtr4	0.000	0.710	-0.737	-0.095	0.990	0.109	-0.083

Table 7: Estimated parameters in the GARCH density specification

6.3.3 Nonparametric density function

The beauty in choosing a non-parametric model is that there is no subjective choice of parameters or restrictive model assumptions. The methodology of [Duan \(2002\)](#); [Stutzer \(1996\)](#); [Arajo and Mar \(2006\)](#) is further beneficial in that the density function is estimated using the only the underlying asset as an input.

Each day the non-parametric density function is estimated. We use 5 years

of historic return data ($5 \times 252 = 1260$ observations) preceding the day of estimation to calculate inputs into our estimation procedure. After calculating the empirical cumulative distribution function and empirical density function, we solve for λ_t to minimise the relative entropy between our function and the standard normal density function of the standardised return; subject to the constraint that the expected return on the distribution is equal to the risk free rate less the dividend yield. We choose the risk free rate equal to 3-month T-Bill rate. The optimal parameter λ_t that minimises relative entropy is shown in table 8 (quarterly average).

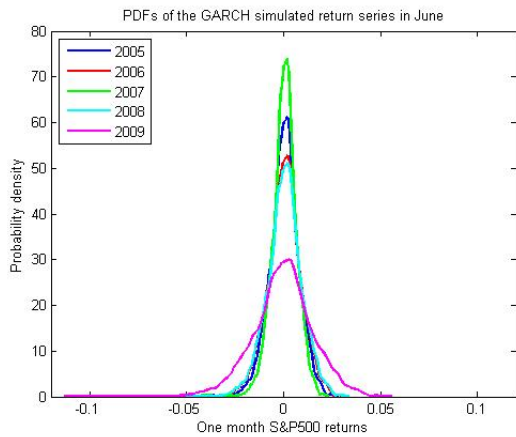
We simulate returns from our density function by sampling normal variables with mean λ_t and variance 1. We then find the cumulative distribution function of this sample. Following which we invert this (simulated) distribution and find each point from the empirical density function using linear interpolation. Our return is then equal to sample estimated standard deviation times this inverse empirical density summed to our sample mean.

	λ_t				
	2005	2006	2007	2008	2009
Qtr1	-0.045	-0.060	-0.100	0.069	0.422
Qtr2	-0.046	-0.078	-0.066	0.108	0.403
Qtr3	-0.030	-0.093	-0.044	0.156	0.368
Qtr4	-0.039	-0.107	-0.016	0.324	0.349

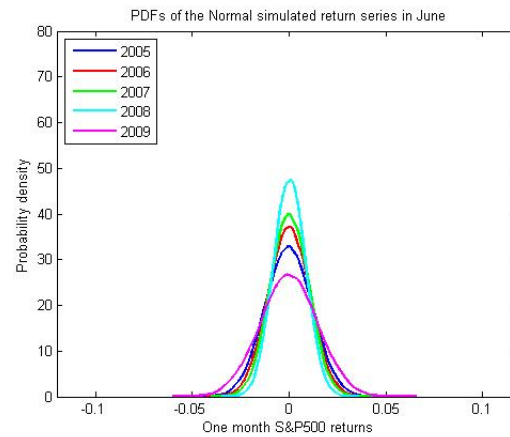
Table 8: Estimated parameters in the Non-parametric density specification

6.3.4 Density function analysis over time

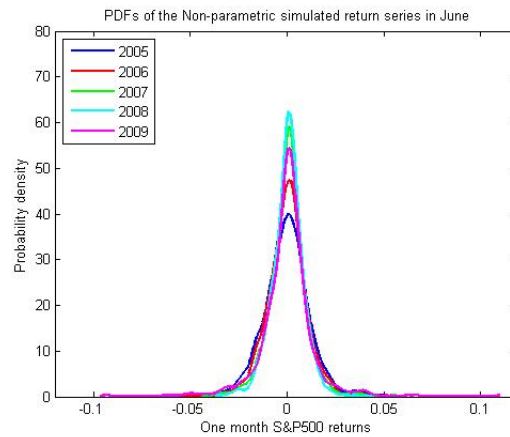
Figure 11 describes our simulated distributions every June each year (details on the simulation are provided in section 6.4). The varying nature of these historic distributions lends itself to a time varying pricing kernel estimate. Additionally, we show probability plots (figure 12) of the simulated returns in June 2009 against a normal distribution and a t -distribution where the parameters for the normal distribution (μ and σ) and the t -distribution ($\hat{\mu}$, $\hat{\sigma}$ and degrees of freedom) are solved through maximum likelihood estimation. This gives some indication as to how the density functions capture the leptokurtic nature of stock returns.



(a) GARCH returns

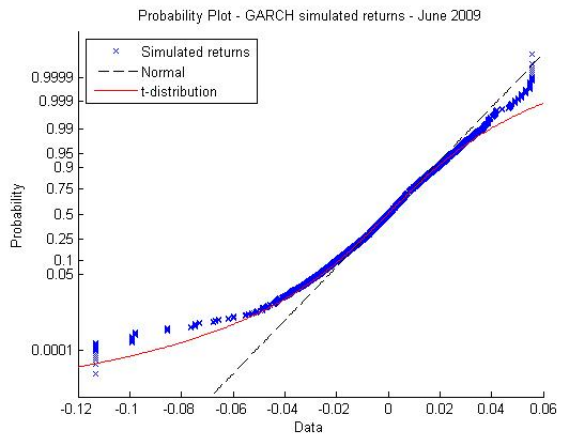


(b) Normal returns

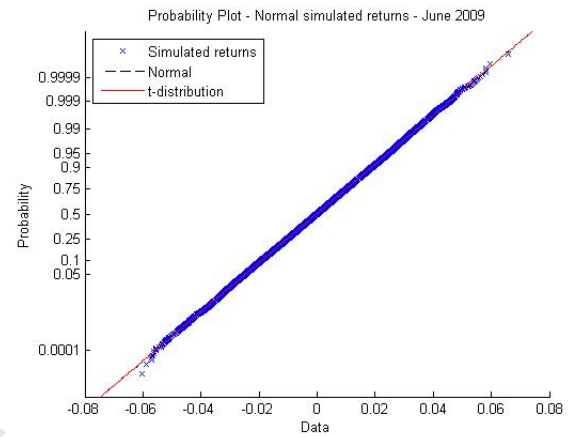


(c) Non-parametric returns

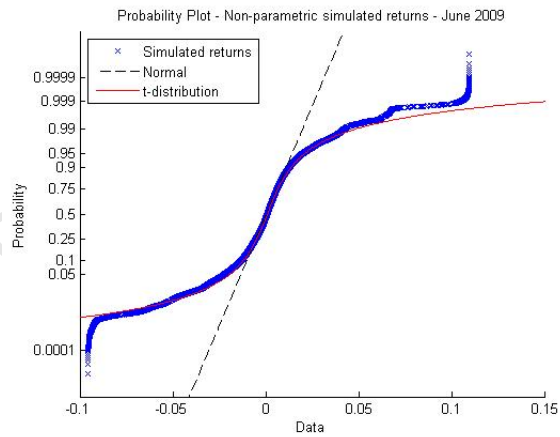
Figure 11: Probability density functions



(a) GARCH returns



(b) Normal returns



(c) Non-parametric returns

Figure 12: Probability plot of returns against Normal and *t*-distributions

6.4 Empirical pricing kernel

We have specified how to calibrate our stochastic volatility models to the market surface and calculate “market” option prices using these models. In addition, we can find our historic densities through Monte Carlo simulation. Now we are in a position to calculate our pricing kernel parameters using the methodology in section 4.3.

Recall that the problem we need to solve to find our asset pricing kernel is equation 41:

$$\min_{\theta_t} \sum_{i=1}^N [p_{i,t} - \widehat{p}_{i,t}(\theta_t)]^2,$$

We find $p_{i,t}$ using our stochastic volatility model to price the option. $\widehat{p}_{i,t}(\theta_t)$ is estimated as follows (equation 42):

$$\begin{aligned} \widehat{p}_{i,t}(\theta_t) &= \frac{1}{H} \sum_{h=1}^H [M_t^*(r_{t+1,h}; \theta_t) g_i(r_{t+1,h})] \\ &= \frac{1}{H} \sum_{h=1}^H [\theta_{0,t}(r_{t+1,h})^{-\theta_{1,t}} g_i(r_{t+1,h})], \end{aligned}$$

We use Monte Carlo simulation to find these model prices. In estimating the pricing kernel, we find our optimal parameters ($\theta_{0,t}$ and $\theta_{1,t}$) on a monthly basis. We restrict calculating market option prices to strikes that range between 90% - 110% moneyness and tenor of one business month (19-23 days). This restriction is to ensure that we price options from our surface where there is least likely a difference compared to the actual market surface. This restriction also reduces our computational time. Our procedure is as follows:

1. Each day we have seven option prices calculated with a fixed tenor but varying strike using our stochastic volatility model. We denote this as our *market* price.
2. Daily, we simulate (80000 simulations) our density functions to find stock returns from which we calculate our model option prices.
3. Our optimisation procedure collates this information on a monthly basis and then finds optimal parameters for our pricing kernel that minimises the difference between the model prices and market prices.

Linking this to our results derived in section 3.3, we summarise our density specifications for returns as follows:

Risk Neutral density	Historic density
Heston	Normal
Bates	GARCH
	Non-parametric

Figures 13 and 14 describe the asset pricing kernel across S&P500 return states. Not only is the time varying nature of the pricing kernel evident, but we are able to see how investor behaviour varies across asset return states. These results are consistent with results achieved by Rosenberg and Engle (2002); Detlefsen et al. (2007). Recall that a decreasing pricing kernel is indicative of an investor's decreasing demand for a unit payoff across states, while an increasing pricing kernel indicates an increasing desire for a unit payoff across states.

6.4.1 Empirical risk aversion

Given the time varying nature of the pricing kernel and our functional form for this kernel, we are able to extract time varying relative risk aversion estimates. Figures 15 and 16 describe parameters $\theta_{0,t}$ and $\theta_{1,t}$. The parameter $\theta_{1,t}$ is our measure of empirical relative risk aversion.

The results achieved under an EGARCH density as the historical density are somewhat surprising as we see a downward trend in risk aversion over the period 2007-2008, namely during the market crash of 2008, a period of high volatility. These results seem contrary to intuition, and contrary to results achieved by Rosenberg and Engle (2002), who also suggest a GARCH specification for the historic density function.

Rosenberg and Engle (2002) show that there is positive correlation between the estimate of risk aversion and the credit spread. They note that the credit spread is a counter cyclical business indicator and hence conclude that risk aversion is also counter cyclical. The empirical findings here and in Rosenberg and Engle (2002) provide support for habit persistence models of investor utility ⁵.

Given that we find our Bates model provides a better fit to the market implied volatility surface, we are inclined to conclude that the the prices of options from the Bates model give us better approximations to market prices. In proceeding, we limit our analysis to the results under the Bates specification as the risk neutral density.

⁵The reader is referred to Rosenberg and Engle (2002); Cochrane (2002); Detlefsen et al. (2007) for further information on different utility functions and their implications for investor behaviour.

We chose to regress our measure of relative risk aversion against certain variables. We regress risk aversion against the one month lagged estimate, the one-month index return, the change in the VIX index and change in the TED spread. Our measure of market volatility is the VIX index and credit risk is measure by the TED spread. Table 9 contains the results.

University of Cape Town

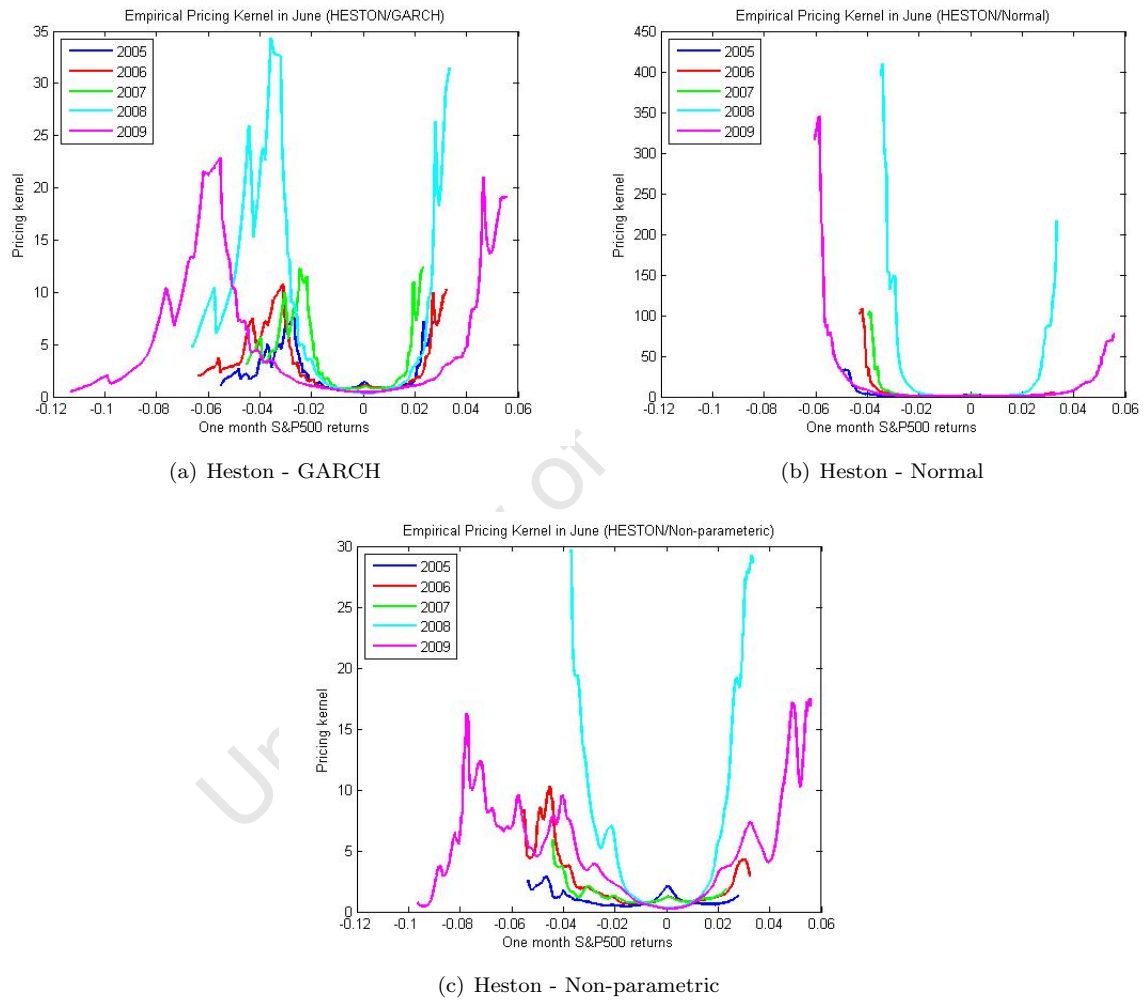


Figure 13: Empirical pricing kernel in June - Heston

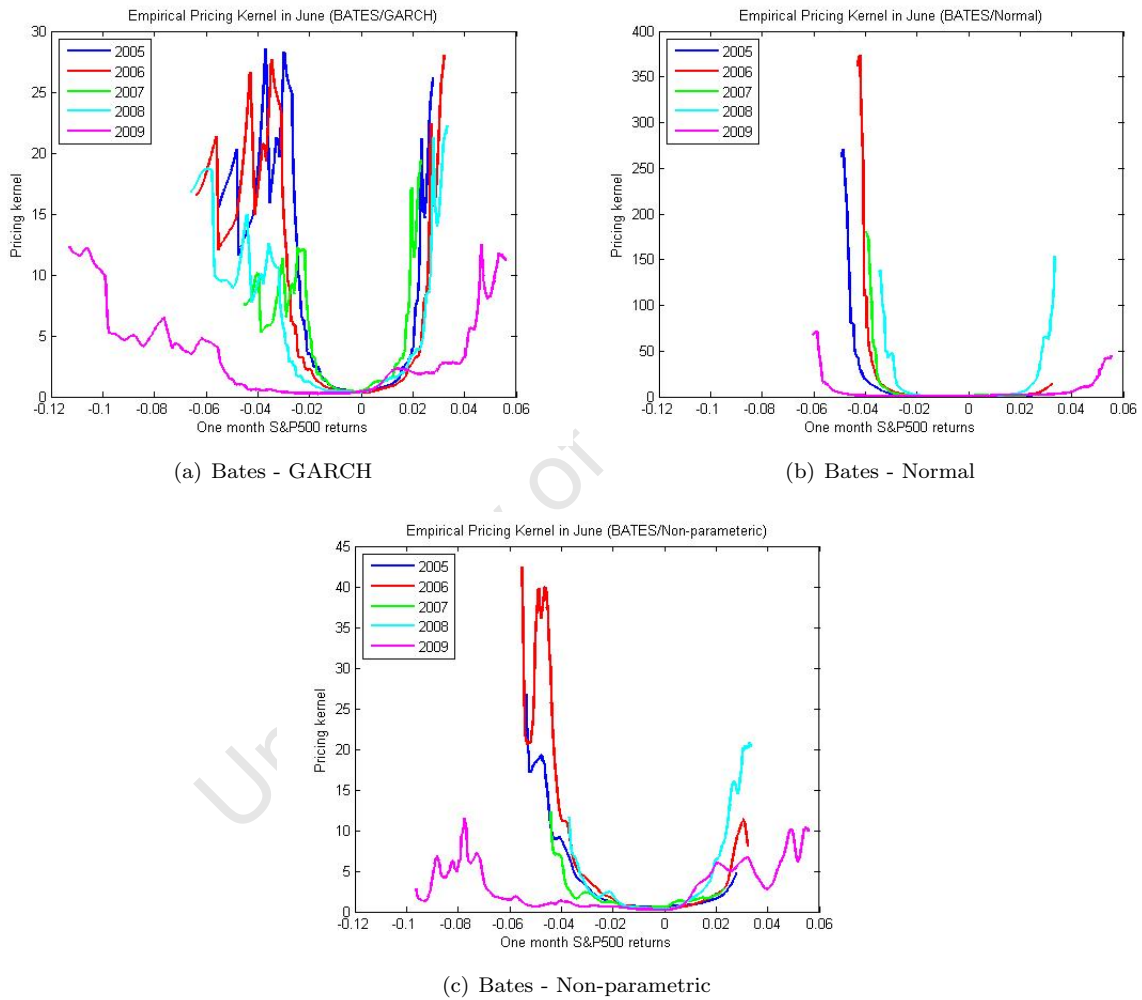


Figure 14: Empirical pricing kernel in June - Bates

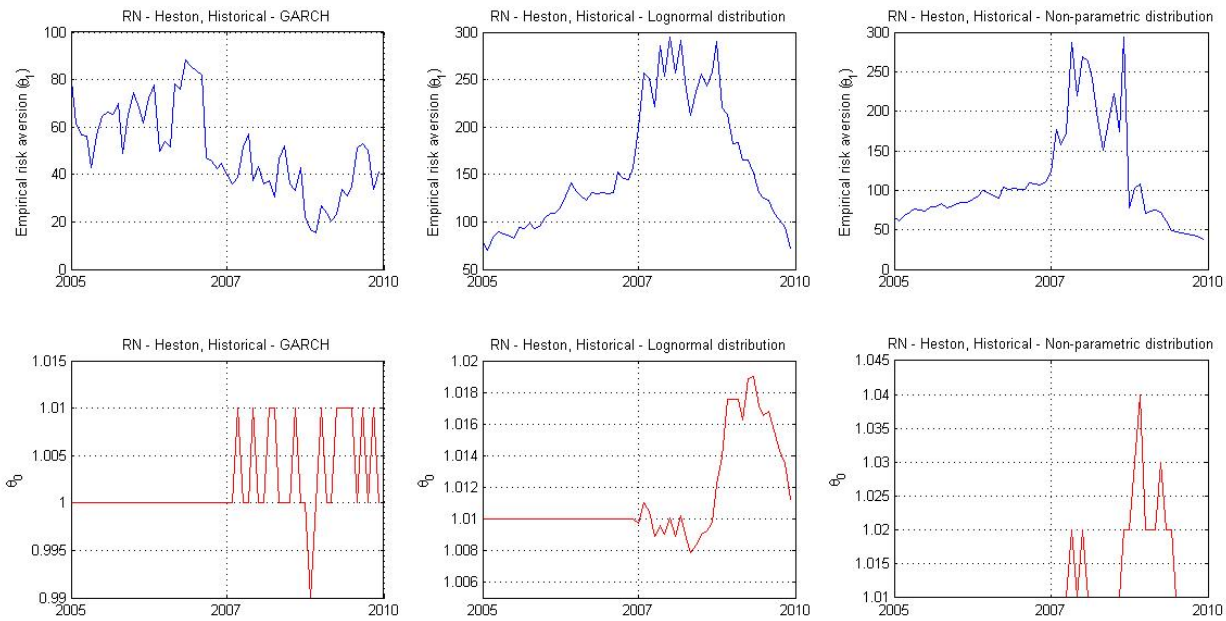


Figure 15: Empirical price kernel parameters - Heston

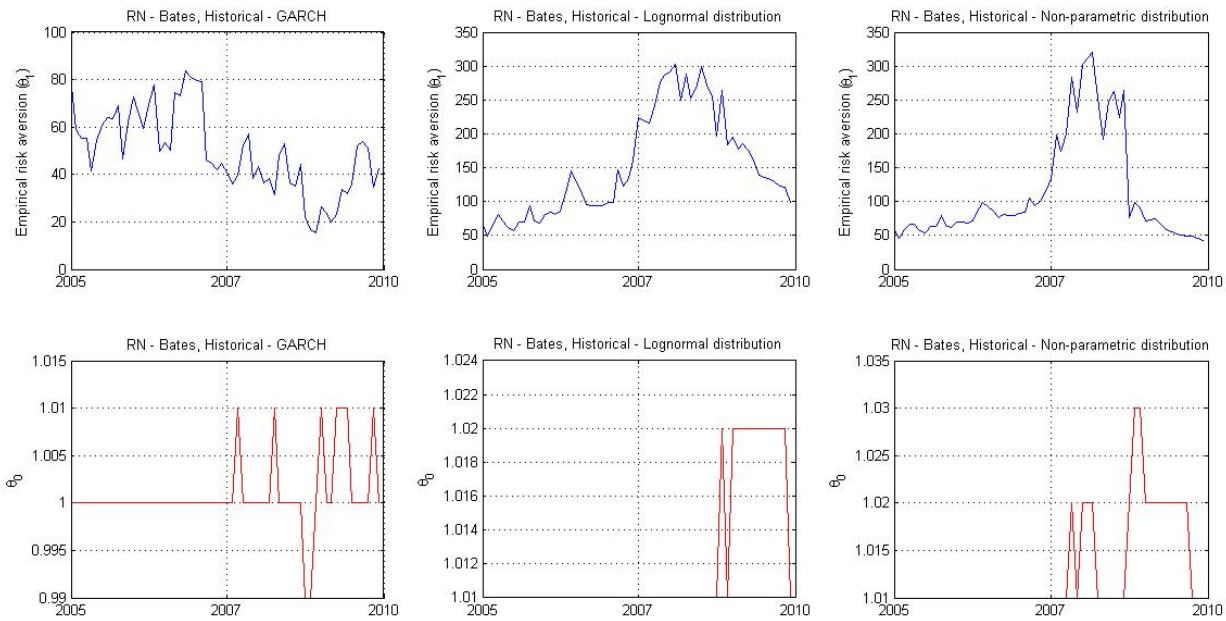


Figure 16: Empirical price kernel parameters - Bates

We find a negative coefficient for the change in the TED spread for all historic density specifications. Initially this result seems to support the [Rosenberg and Engle \(2002\)](#) theory as it suggests that risk aversion is negatively correlated to credit risk (implying positive correlation to the credit spread) and hence that risk aversion is counter cyclical; however, we find that this coefficient estimate is only significant for the normal density specification.

Our coefficient estimates for the one-month lagged risk aversion estimate and the VIX index in the normal and non-parametric specifications seems more intuitive. We see that risk aversion is a function of historical risk aversion as well as a function of market volatility. Assuming that market volatility is an indicator of risk appetite and the overall economy, we can deduce that risk aversion is counter cyclical, a result consistent with the findings of [Rosenberg and Engle \(2002\)](#) and the theory of [Fama and French \(1989\)](#). If we assume an investor's wealth is increasing over time, an increase in relative risk aversion suggests a smaller holding in risky assets and vice versa. We aim to show this result by including the underlying index in figure 17.

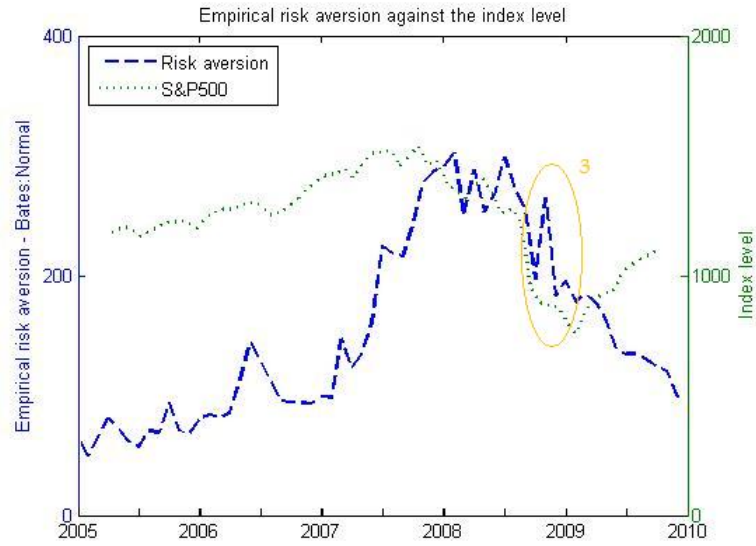
Understanding the evolution of risk aversion and asset prices provides a useful foundation in understanding investor behaviour. Risk aversion estimates could also provide predictive power to understand the progression of the underlying index. We choose three events (highlighted in figure 17) to narrate our estimates for the Bates/Normal and Bates/Non-parametric risk aversion estimates. This is clear evidence in support that investor risk aversion is counter cyclical.

1. 26 July 2007 ([Dennis, 2007](#))
Investors are exhibiting increased aversion to risk in equity and currency markets fearing a dry up of deals in the leverage buy out market as well as increasing concerns about poor home sales data in the housing market. The US data released on the housing market showed that existing home sales fell to the lowest in five years. Commodities markets were lower indicating caution around a slowdown in the United States, while credit markets reported significant bad debts.
2. 2 July 2008 ([Garnham, 2008](#))
Strong economic figures in Japan drove the yen higher. Economic data came out stronger than expected which abated investor fear. Traders however, were still cognisant of a slowdown in equity markets in United Kingdom and in the United States due to falling house prices.
3. 6 September 2008 ([Oakley and Mackenzie, 2008](#))
Financial markets tumbled on concerns of a distraught housing market and the wellbeing of banks. As the US Treasury finalised a deal to increase the capital at mortgage firms, Fannie Mae and Freddie Mac, commodity

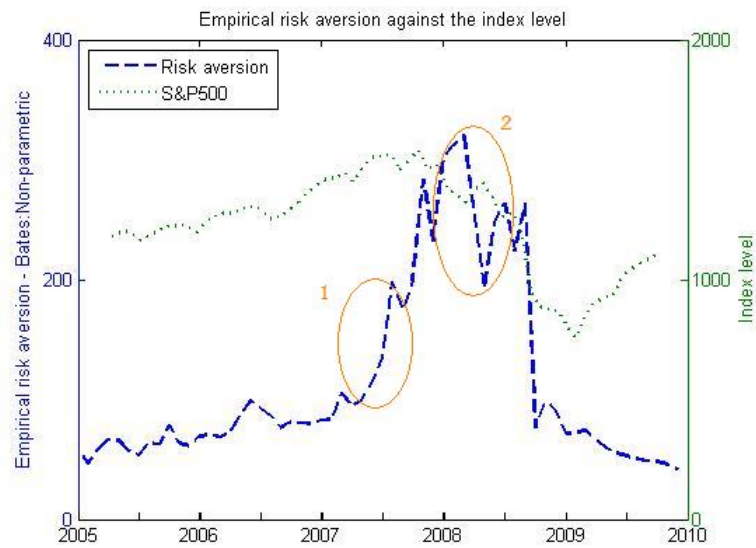
prices fell and the S&P500 lost 3.2%. Increased risk aversion saw investors moving into safer havens such as 10-year bonds.

BATES_GARCH			
	Coefficient	t-stat	p-value
Constant	12.52	2.84	0.01
One month lagged RRA	0.74	8.75	0.00
S&P500 monthly return	143.02	2.70	0.01
Change in VIX	0.97	1.70	0.09
Change in TED	-0.07	-1.39	0.17
R Squared	69%		
Adjusted R Squared	66%		
BATES_Normal			
	Coefficient	t-stat	p-value
Constant	7.27	1.09	0.28
One month lagged RRA	0.96	24.22	0.00
S&P500 monthly return	20.97	0.19	0.85
Change in VIX	6.23	5.47	0.00
Change in TED	-0.51	-5.01	0.00
R Squared	91%		
Adjusted R Squared	90%		
BATES_Non-parametric			
	Coefficient	t-stat	p-value
Constant	1.25	0.26	0.80
One month lagged RRA	1.02	28.57	0.00
S&P500 monthly return	-12.36	-0.12	0.90
Change in VIX	4.62	4.44	0.00
Change in TED	-0.12	-1.24	0.22
R Squared	82%		
Adjusted R Squared	80%		

Table 9: Regression results



(a) Bates - Normal



(b) Bates - Non-parametric

Figure 17: Empirical risk aversion

7 Conclusion

Explaining investor sentiment and behaviour is a complex sphere of economic research. This paper has aimed to make a useful inroad in understanding investor risk aversion in the US market over the period 2005 to 2009, namely during the equity market crash of 2008. Importantly, the methodology suggested here provides consistent results that investor risk aversion is counter cyclical. We suggest that a stochastic volatility model with jumps and a non-parametric model for the risk neutral and historic density functions respectively be a tractable methodology for emerging markets or illiquid markets.

We calibrated our stochastic volatility models to implied volatility data and enjoyed success in that the Bates model is able to capture salient features of the market surface. We feel that our assumptions and optimisation techniques can be defended in respect of this methodology. From our stochastic volatility model, we are able to extract a risk neutral density function. Our historic density specifications are calculated from an EGARCH, normal and non-parametric model for returns. Being able to collate this information on density functions, we make no prior assumptions on the investor utility function other than the specification of the function form of the SDF.

Most importantly, we follow a simple power functional form of the asset-pricing kernel, which is a function of returns. We find evidence of a time varying pricing kernel across return states. Furthermore, we are able to extract risk aversion estimates from this pricing kernel. Our results from a normal and non-parametric historic density function indicate that investor's relative risk aversion increased over the period 2007-2008 during the market crash. These results are consistent with investor behaviour presented in market reports and the state of the overall economy at the time.

Our estimates, particularly from our non-parametric model of returns provide a useful toolbox for investors in emerging markets where derivative instruments are not liquidly traded. If there is sufficient information to calibrate a stochastic volatility model (using proxies) an investor is now in a position to make useful judgments on the extent of risk aversion in the market.

Our generalised asset-pricing framework can be improved with the use of market prices of derivative securities, however the merit of this methodology is that in illiquid markets there is now a method to extract risk aversion estimates consistent with the market. Obviously, opportunity exists in researching the extent to which proxies can be used to find a volatility surface in an illiquid market and what the range of proxies may even be.

Having risk aversion estimates consistent with intuition and the economy, suggests that our functional form for the pricing kernel may be useful in creating a trading strategy to take advantage of possible mispricing in a derivatives mar-

ket. [Rosenberg and Engle \(2002\)](#) indicate the superior hedging performance of this functional form. The opportunity exists to test these results in emerging and illiquid markets.

Given our initial objective, we have successfully postulated a methodology to find the state price density and historical density function and use this to reveal characteristics of the stochastic discount factor and consequently risk aversion of an investor over time.

University of Cape Town

A Mathematical derivations

A.1 Deriving the empirical risk aversion parameter

Recall that the measure of relative risk aversion using the projected pricing kernel is:

$$\gamma_t^* = -\frac{X_{t+1}M_t'^*(X_{t+1})}{M_t^*(X_{t+1})}$$

Assuming the functional form for the pricing kernel as in [Rosenberg and Engle \(2002\)](#):

$$M_t^*(r_{t+1}) = \theta_{0,t}(r_{t+1})^{-\theta_{1,t}}$$

Taking the first derivative:

$$M_t^{*'}(r_{t+1}) = \theta_{0,t}(-\theta_{1,t})(r_{t+1})^{-\theta_{1,t}-1}$$

Which implies that:

$$\begin{aligned}\gamma_t^* &= \frac{r_{t+1}\theta_{0,t}(-\theta_{1,t})(r_{t+1})^{-\theta_{1,t}-1}}{\theta_{0,t}(r_{t+1})^{-\theta_{1,t}}} \\ &= \theta_{1,t}\end{aligned}$$

A.2 Deriving the distribution of stock prices

A.2.1 Ito's formula

Let S_t be a stochastic integral:

$$dS_t = adt + bdW_t$$

and $g(t, s)$ be a twice differential function on $[0, T] \times \mathbb{R}$.

Then the stochastic process $Y_t = g(t, S_t)$ is a stochastic integral with

$$dY_t = \frac{\delta g}{\delta t}(t, S_t)dt + \frac{\delta g}{\delta s}dS_t + \frac{1}{2}\frac{\delta^2 g}{\delta s^2}(t, S_t)(dX_t)^2$$

A.2.2 Derivation

The Geometric Brownian Motion of stock prices:

$$\frac{dX_t}{X_t} = \mu dt + \sigma dW_t$$

To solve this stochastic differential equation we proceed using Ito's formula with $g(t, y) = \ln y$.

Then:

$$\begin{aligned}\frac{\delta g}{\delta t} &= 0 \\ \frac{\delta g}{\delta x} &= \frac{1}{x} \\ \frac{\delta^2 g}{\delta x^2} &= -\frac{1}{x^2}\end{aligned}$$

Using Ito's formula we obtain:

$$\begin{aligned}d(\ln X_t) &= \frac{1}{X_t} dX_t - \frac{1}{2} \frac{1}{X_t^2} (dX_t)^2 \\ &= \frac{dX_t}{X_t} - \frac{1}{2} \frac{1}{X_t^2} (\mu X_t dt + \sigma X_t dW_t)^2 \\ &= \frac{dX_t}{X_t} - \frac{1}{2} (\mu dt + \sigma dW_t)^2 \\ &= \frac{dX_t}{X_t} - \frac{1}{2} \sigma^2 dt \\ \mu dt + \sigma dW_t - \frac{1}{2} \sigma^2 dt &= \mu - \frac{1}{2} \sigma^2 dt + \sigma dW_t\end{aligned}$$

Integrate:

$$\ln X_t = \ln X_0 + \int_0^t (\mu - \frac{1}{2} \sigma^2) du + \int_0^t \sigma dW_u$$

We obtain the log normal evolution of stock prices:

$$\begin{aligned}\ln \frac{X_t}{X_0} &= (\mu - \frac{1}{2} \sigma^2) t + \sigma W_t \\ X_t &= X_0 e^{(\mu - \frac{1}{2} \sigma^2) t + \sigma W_t}\end{aligned}$$

If we let $r_X = \ln \frac{X_t}{X_0}$ (the log return on X), we can follow the properties of a Brownian Motion W_t to derive that r_X is normally distributed.

$$W_t \sim N(0, T - t)$$

Following the properties of a normal distribution, it is easy to see that:

$$r_X = \ln \left(\frac{X_t}{X_0} \right) \sim N \left(\left(\mu - \frac{\sigma^2}{2} \right) (T - t), \sigma^2 (T - t) \right)$$

B Mathematical proofs

B.1 Risk neutral dynamics of a jump process

Here we show that the risk neutral dynamics proposed in section 4.1.1 for the asset price process in the Bates (1996) model is a martingale.

Consider the following stochastic differential equation for the evolution of the stock price process (Sepp, 2003):

$$dX_t = \mu(X, t)dt + \sigma(X, t)dW_t + \gamma(X, t)JdN_t$$

where N_t is a Poisson process with constant intensity λ . The jump magnitude J has density function $\bar{w}(J)$.

These risk neutral dynamics are chosen in such a way that the discounted asset price process remains a martingale. This proof is based on Sepp (2003).

To calculate the expected value of the increment dN_t , we follow the properties of the Poisson process that says that the arrival of the next jump is independent of the arrival of previous jumps and that the probability of two simultaneous jumps is zero. Assume that j jumps have occurred. Over the next time interval dt :

$$\begin{aligned} \mathbb{P}[\text{one jump occurring}] &= \lambda dt \\ \mathbb{P}[\text{no jumps occurring}] &= 1 - \lambda dt \end{aligned}$$

Therefore

$$\mathbb{E}[dN_t] = 1 \times \lambda dt + 0 \times (1 - \lambda dt) = \lambda dt$$

Define the process $dM_t = dN_t - \lambda dt$; we can clearly see that $\mathbb{E}[dM_t] = 0$ and thus dM_t is a martingale. Given that we have specified our density function for our jump size as a normal density with mean v and standard deviation δ , our risk neutral asset dynamics are:

$$dX_t = (\mu - \lambda m)X_t dt + \sigma X_t dW^s + (e^J - 1)X_t dN_t$$

where m is the average jump amplitude: $m = e^{v + \frac{1}{2\delta^2}} - 1$.

B.2 Relative entropy

The relative entropy between a distribution f and a distribution g is a measure of the distance between the two distributions. The relative entropy is always

non-negative and equal to zero if the distribution f is exactly equal to the distribution g .

If we follow the definition of [Derman and Zou \(1999\)](#), the decrease in entropy between two distributions f and g is measured as follows:

$$S(f, g) = \mathbb{E}_g[\log(g) - \log(f)] = \int_x g(x) \log\left(\frac{g(x)}{f(x)}\right) dx$$

In section [4.2.3](#) we indicate that using the relative entropy principle, the density function for the normalised asset return is given by minimising the following integral ([Duan, 2002](#)):

$$\int_{-\infty}^{\infty} f(x) \ln\left(\frac{f(x)}{\phi(x)}\right) dx$$

subject to the constraints:

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) dx &= 1 \\ \int_{-\infty}^{\infty} x f(x) dx &= a_t \end{aligned}$$

where:

a_t | Constant that is a function of time

We state that the solution is the density function defined as follows ([Cover and Thomas, 1991](#)):

$$f(x, \lambda_t) = \frac{\phi(x)e^{\lambda_t x}}{\int_{-\infty}^{\infty} \phi(x)e^{\lambda_t x} dx} = \phi(x - \lambda_t)$$

We include the proof by [Buchen and Kelly \(1996\)](#) for sake of completion.

Proof. The proof requires the use of Lagrange multipliers:

Let

$$S(f, g) = \int_{-\infty}^{\infty} g(x) \ln\left(\frac{g(x)}{f(x)}\right) - (1 + \lambda_0) \int_{-\infty}^{\infty} g(x) dx - \sum_{i=1}^m \lambda_i \int_{-\infty}^{\infty} x g(x) dx$$

Take the derivative w.r.t $g(x)$ (applying the product rule):

$$\delta S = \left[\int_{-\infty}^{\infty} \ln\left(\frac{g(x)}{f(x)}\right) dx + \int_{-\infty}^{\infty} g(x) \frac{f(x)}{g(x)} \frac{1}{f(x)} dx - \sum_{i=1}^m \lambda_i \int_{-\infty}^{\infty} x dx \right] \delta g(x)$$

This reduces to:

$$\delta S = \int_{-\infty}^{\infty} \left[\ln \frac{g(x)}{f(x)} - \lambda_0 - \sum_{i=1}^m \lambda_i x \right] \delta g(x) \, dx$$

Solving for $g(x)$ yields:

$$g(x) = \frac{f(x)}{\Psi} e^{\sum_{i=1}^m \lambda_i x}$$

where

$$\Psi = \int_{-\infty}^{\infty} e^{\sum_{i=1}^m \lambda_i x} \, dx$$

Now, set

$$\sum_{i=1}^m \lambda_i = \lambda_t$$

and we have the result as required:

$$g(x, \lambda_t) = \frac{f(x) e^{\lambda_t x}}{\int_{-\infty}^{\infty} f(x) e^{\lambda_t x} \, dx} = f(x - \lambda_t) \quad \square$$

C Option pricing under affine functions

C.1 Characteristic function

Given the SDE :

$$dX = \mu(X, t)dt + \sigma(X, t)dW$$

An Affine system is developed where the expectations of exponentials can always be written in the form:

$$\phi(\mu, X_t, \tau) = e^{A(\mu, \tau) + B(\mu, \tau)^T X_t}$$

This solutions to this system are given explicitly below:

$$\begin{aligned} B_1(u, \tau) &= iu_1 \\ B_2(u, \tau) &= \frac{(\beta - D) - Ge^{-D\tau}(\beta + D)}{\omega^2(1 - Ge^{-D\tau})} \\ A(u, \tau) &= \frac{\kappa\theta}{\omega^2} \left((\beta - D) - 2 \log \left(\frac{Ge^{-D\tau} - 1}{G - 1} \right) \right) \\ G(u) &= \frac{\beta - D - iu_2\omega^2}{\beta + D - iu_2\omega^2} \\ D(u) &= \sqrt{\beta^2 - 4\alpha\gamma} \\ \alpha &= -\frac{1}{2}u_1(i + u_1) \\ \beta &= \kappa - \rho\omega u_1 i \\ \gamma &= \frac{1}{2}\omega^2 \end{aligned}$$

Taking the initial condition to be at $t = 0$, we obtain:

$$\phi(u, (x, v), 0) = e^{A(u, \tau) + B_1(u, \tau)x_0 + B_2(u, \tau)v_0}$$

Using the characteristic function of $\log S_T = x_T + rT$, and also that $x_0 = \log S_0 = 0$. We obtain [Becker \(2009\)](#):

$$\tilde{\phi}(u) = e^{iur\tau + A(u, \tau) + B_2(u, \tau)v_0}$$

C.2 Fourier transforms

A *Fourier Transform* of a function $f : \mathbb{R} \rightarrow \mathbb{R}$ (FT(f) = $\mathfrak{F}(f) = \hat{f}$) is defined as:

$$\hat{f}(u) = \int e^{iux} f(x) dx$$

and the *Inverse Fourier Transform* of a function $g : \mathbb{R} \rightarrow \mathbb{R}$ is defined as:

$$\mathfrak{F}^{-1}(g)(x) = \frac{1}{2\pi} \int e^{-iux} g(u) du$$

Hence:

$$\mathfrak{F}^{-1}(\hat{f})(x) = f(x)$$

Let f be a density function, and let ϕ be it's characteristic function, then [Becker \(2009\)](#):

$$\int_x^\infty f(t) dt = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{1}{u} \mathfrak{F}(e^{-iux} \phi(u)) du$$

C.3 Computing the Fourier integral

The Fourier integral may be computed by means of standard procedures of numerical integration. To approximate the Fourier integral it is rewritten:

$$I(k) = \int_0^\infty f(k) dk = \lim_{N \rightarrow \infty} \sum_{j=1}^N \int_{(j-1)h}^{j.k} f(k) dk \quad (43)$$

Where the sub-quadrature are computed by means of Gaussian Quadrature formula ([Sepp, 2003](#)). The N value is fixed to a relatively large value, in the MATLAB code it is fixed to 100, which corresponds to $K = \log 100$, deemed a sufficiently large value. The `quadl` function is used in MATLAB to compute the Fourier integral.

D Solving nonlinear least squares problems

We use the function `lsqnonlin` in MATLAB. This function solves problems of the form:

$$\min_x \|f(x)\|^2 = \min_x [f_1(x)^2 + f_2(x)^2 + \dots + f_n(x)^2]$$

$x = \text{lsqnonlin}(\mathbf{fun}, x_0, \mathbf{lb}, \mathbf{ub})$ starts at a point x_0 and finds a minimum of the sum of the squares of the functions described in `fun`. Furthermore, a set of lower and upper bounds are defined by `lb` and `ub` respectively that ensure $\mathbf{lb} \leq x_0 \leq \mathbf{ub}$. This is a local optimizer as the choice of x_0 is subjective and the solution is based on this initial choice.

The problem in this paper is a large scale problem meaning that the solution space is bounded and that the system is not underdetermined. Given our parameter bounds (section 6.2) and the number of options we have to calibrate our surfaces, our problem is large scale. This type of problem is solved by a trust-region-reflective algorithm. The interested reader is referred to the references from MATLAB, namely [Coleman and Li \(1996, 1994\)](#).

E Simulation of probability densities

In section 4.1.1 we define the continuous version of the [Heston \(1993\)](#) model:

$$\begin{aligned} \frac{dX_t}{X_t} &= \mu dt + \sqrt{V_t} dW_t^s \\ dV_t &= \kappa(\theta - V_t) dt + \sigma\sqrt{V_t} dW_t^q \end{aligned}$$

where:

X_t	Underlying asset price
V_t	Variance
W_t^s, W_t^q	Standard Brownian Motions
κ	Mean reversion parameter
θ	Long run variance
σ	Volatility of variance
ρ	Correlation parameter

Brownian motion processes W_t^s and W_t^q are correlated with parameter ρ such that $dW^s \cdot dW^q = \rho dt$.

To simulate returns or stock prices from this stochastic process, we discretise the stochastic process using the Euler-Maruyama method. This yields:

$$\begin{aligned} X_t &= X_{t-1} + \mu X_{t-1} dt + \sqrt{V_{t-1}} X_{t-1} \sqrt{dt} W_t^s \\ V_t &= V_{t-1} + \kappa(\theta - V_{t-1}) dt + \sigma \sqrt{V_{t-1}} \sqrt{dt} W_t^q \end{aligned}$$

where W_t^s and W_t^q are standard normal random variables with correlation ρ .

We use the Cholesky decomposition to make these variables functions of *independent* standard normal random variables (Becker, 2009):

$$\begin{aligned} W_t^s &= v_t^s \\ W_t^q &= \rho v_t^s + \sqrt{1 - \rho^2} v_t^q \end{aligned}$$

where v_t^s and v_t^q are independent standard normal random variables.

In section 4.1.2 we define the continuous version of the Bates (1996) model:

$$\begin{aligned} dX_t &= (\mu - \lambda m)X_t dt + \sqrt{V_t}X_t dW^s + (e^J - 1)X_t dN_t \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW^q \end{aligned}$$

where:

$N(t)$	A Poisson process with constant intensity λ
J	Jump magnitude (random variable) with density function $\bar{w}(J)$
m	$e^{v + \frac{1}{2\delta^2}} - 1$

We discretise this stochastic process to yield:

$$\begin{aligned} X_t &= (\mu - \lambda m)X_{t-1}dt + \sqrt{V_{t-1}}X_{t-1}\sqrt{dt}W_t^s + (e^J - 1)X_{t-1}N_t \\ V_t &= V_{t-1} + \kappa(\theta - V_{t-1})dt + \sigma\sqrt{V_{t-1}}\sqrt{dt}W_t^q \end{aligned}$$

where:

W_t^s and W_t^q are standard normal random variables with correlation ρ .
 N_t is a Poisson counter drawn from a Bernoulli distribution with parameter λ .
 J is sampled from a normal distribution with mean v and standard deviation δ .

References

- Yacine Ait-Sahalia and Andrew W. Lo. Nonparametric estimation of state-price densities implicit in financial asset prices. *The Journal of Finance*, 53(2):499–547, 1998. ISSN 00221082. URL <http://www.jstor.org/stable/117360>.
- Yacine Ait-Sahalia and Andrew W. Lo. Nonparametric risk management and implied risk aversion. *Journal of Econometrics*, 94(1-2):9 – 51, 2000. ISSN 0304-4076. URL <http://www.sciencedirect.com/science/article/B6VC0-40B2MF9-2/2/9db6778e229b445db56f459183497381>.
- Mark De Arajo and Eben Mar. Examining the volatility skew in the south african equity market using risk-neutral historical distributions. *Investment Analysts Journal*, 2006.
- Kenneth Arrow. *Essays in the Theory of Risk-Bearing*. Chicago: Markham, 1971.
- Gurdip Bakshi, Charles Cao, and Zhiwu Chen. Empirical performance of alternative option pricing models. *The Journal of Finance*, 52(5):2003–2049, 1997. ISSN 00221082. URL <http://www.jstor.org/stable/2329472>.
- David S. Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *The Review of Financial Studies*, 9(1):69–107, 1996. ISSN 08939454. URL <http://www.jstor.org/stable/2962366>.
- Ronald Becker. Stochastic calculus ii lecture notes. Local, Implied and Stochastic Volatility Models, 2009.
- Lorenzo Bergomi. Smile dynamics 2. *Risk Magazine*, pages 67–73, 2005. URL <http://ssrn.com/abstract=1520443>.
- Thomas Bjork. *Arbitrage Theory in Continuous Time*. Oxford University Press, second edition, 2004.
- Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):pp. 637–654, 1973. ISSN 00223808. URL <http://www.jstor.org/stable/1831029>.
- Peter W. Buchen and Michael Kelly. The maximum entropy distribution of an asset inferred from option prices. *The Journal of Financial and Quantitative Analysis*, 31(1):pp. 143–159, 1996. ISSN 00221090. URL <http://www.jstor.org/stable/2331391>.
- CBOE. Introduction to vix futures and options. Chicago Board Options Exchange, 2010. URL www.cboe.com/micro/vix/introduction.aspx.
- Carl Chiarella, Mark Craddock, and Nadima El-Hassan. The calibration of stock option pricing models using inverse problem methodology. *Research Paper Series: QFRC*, 39, 2000.

- Peter Christoffersen, Steven Heston, and Kris Jacobs. Option valuation with conditional skewness. *Journal of Econometrics*, 131(1-2):253 – 284, 2003. ISSN 0304-4076. URL <http://www.sciencedirect.com/science/article/B6VC0-4FJXNGG-1/2/e701085b7912439e3378a14fb478af21>.
- John H. Cochrane. *Asset Pricing*. Princeton University Press, Princeton and Oxford, 2002.
- Thomas F. Coleman and Yuying Li. On the convergence of reflective newton methods for large-scale nonlinear minimization subject to bounds. *Mathematical Programming*, 67(2):189–224, 1994.
- Thomas F. Coleman and Yuying Li. An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6:418–445, 1996.
- George M Constantinides, Milton Harris, and Rene M Stulz. *Handbook in Economics*. Elsevier, North Holland, 1b edition, 2003.
- Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley, New York, 1991.
- John C. Cox and Stephen A. Ross. The valuation of options for alternative stochastic processes. *Journal of Financial Economics*, 3(1-2):145–166, 1976. URL <http://ideas.repec.org/a/eee/jfinec/v3y1976i1-2p145-166.html>.
- Neil Dennis. Risk aversion threat to buy-outs rattles stocks. *Financial Times*, 2007. URL <http://www.ft.com/cms/s/0/a90c9d76-3b27-11dc-8f9e-0000779fd2ac.html#axzz15F0dyPXV>.
- Emanuel Derman and Iraji Kani. Riding on a smile. *RISK*, 7(2):32–39, February 1994.
- Emanuel Derman and Joseph Zou. Strike adjusted spread: A new metric for estimating the value of equity options. Technical report, Goldman Sachs, 1999.
- Kai Detlefsen, Wolfgang Hardle, and Rouslan Moro. Empirical pricing kernels and investor preferences. *SFB 649 Discussion Paper 2007-017*, 2007.
- Zhuanxin Ding and Clive W. J. Granger. Modeling volatility persistence of speculative returns: A new approach. *Journal of Econometrics*, 73(1):185–215, July 1996. URL <http://ideas.repec.org/a/eee/econom/v73y1996i1p185-215.html>.
- Jin-Chuan Duan. Nonparametric option pricing by transformation. 2002. URL <http://www.rotman.utoronto.ca/jcduan>.

- Robert F. Engle and Chowdhury Mustafa. Implied arch models from options prices. *Journal of Econometrics*, 52(1-2):289–311, 1992. URL <http://ideas.repec.org/a/eee/econom/v52y1992i1-2p289-311.html>.
- Eugene F. Fama and Kenneth R. French. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, 25(1):23 – 49, 1989. ISSN 0304-405X. URL <http://www.sciencedirect.com/science/article/B6VBX-45NHW7H-4/2/aac51e1f952bf528403fa73f3344d06c>.
- Peter Garnham. Yen rises on strong economic figures. *Financial Times*, 2008. URL <http://www.ft.com/cms/s/0/e558be26-47cf-11dd-93ca-000077b07658.html#axzz15FPFklgn>.
- J. Michael Harrison and David M. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3):381–408, June 1979. URL <http://ideas.repec.org/a/eee/jetheo/v20y1979i3p381-408.html>.
- Steven L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993. ISSN 08939454. URL <http://www.jstor.org/stable/2962057>.
- Steven L. Heston and Saikat Nandi. A closed-form garch option valuation model. *The Review of Financial Studies*, 13(3):585–625, 2000. ISSN 08939454. URL <http://www.jstor.org/stable/2645997>.
- John C. Hull. *Options, Futures and other Derivatives*. Pearson Prentice Hall, sixth edition, 2006.
- Jens Carsten Jackwerth. Recovering risk aversion from option prices and realized returns. *The Review of Financial Studies*, 13(2):433–451, 2000. ISSN 08939454. URL <http://www.jstor.org/stable/2646032>.
- Jens Carsten Jackwerth and Mark Rubinstein. Recovering probability distributions from option prices. *The Journal of Finance*, 51(5):1611–1631, 1996. ISSN 00221082. URL <http://www.jstor.org/stable/2329531>.
- Christopher S. Jones. The dynamics of stochastic volatility: evidence from underlying and options markets. *Journal of Econometrics*, 116(1-2):181 – 224, 2003. ISSN 0304-4076. URL <http://www.sciencedirect.com/science/article/B6VC0-48945H4-6/2/ae16c2a350346cd4be21dcd0a36f3aeb>. Frontiers of financial econometrics and financial engineering.
- Yueh-Neng Lin, Norman Strong, and Xinzhong Xu. Pricing ftse 100 index options under stochastic volatility. *Journal of Futures Markets*, 21:197–211, 2001.

- Robert E. Lucas. Asset prices in an exchange economy. *Econometrica*, 46(6): pp. 1429–1445, 1978. ISSN 00129682. URL <http://www.jstor.org/stable/1913837>.
- Robert C. Merton. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3:125–144, 1976.
- Sergei Mikhailov and Ulrich Nogel. Hestons stochastic volatility model implementation, calibration and some extensions. *Wilmott magazine*, 2003.
- Nimalin Moodley. The heston model: A practical approach with matlab code. Master’s thesis, University of the Witwatersrand, 2005.
- Saikat Nandi. How important is the correlation between returns and volatility in a stochastic volatility model? empirical evidence from pricing and hedging in the s&p 500 index options market. *Journal of Banking & Finance*, 22(5):589 – 610, 1998. ISSN 0378-4266. URL <http://www.sciencedirect.com/science/article/B6VCY-3T3SMJY-5/2/7653117287878f65298201da693f5f39>.
- David Oakley and Michael Mackenzie. Investors take cover as bears gain upper hand. *Financial Times*, 2008. URL <http://www.ft.com/cms/s/0/7d38da7e-7baa-11dd-b839-000077b07658.html#axzz15FQFH5gE>.
- Jun Pan. The jump-risk premia implicit in options: evidence from an integrated time-series study. *Journal of Financial Economics*, 63(1):3 – 50, 2002. ISSN 0304-405X. URL <http://www.sciencedirect.com/science/article/B6VBX-44D4PBW-1/2/e201da1cad84a2f0021a9a9ce25b77fe>.
- John W. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32(1/2):122–136, 1964. ISSN 00129682. URL <http://www.jstor.org/stable/1913738>.
- Joshua V. Rosenberg and Robert F. Engle. Empirical pricing kernels. *Journal of Financial Economics*, 64(3):341 – 372, 2002. ISSN 0304-405X. URL <http://www.sciencedirect.com/science/article/B6VBX-45W372J-1/2/7ad0f8d6d071d78e79f611a1a21402f8>.
- Stephen A. Ross. Options and efficiency. *The Quarterly Journal of Economics*, 90(1):75–89, 1976. ISSN 00335533. URL <http://www.jstor.org/stable/1886087>.
- Mark Rubinstein. Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active cboe option classes from august 23, 1976 through august 31, 1978. *The Journal of Finance*, 40(2):455–480, 1985. ISSN 00221082. URL <http://www.jstor.org/stable/2327895>.
- Mark Rubinstein. Implied binomial trees. *The Journal of Finance*, 49(3):771–818, 1994. ISSN 00221082. URL <http://www.jstor.org/stable/2329207>.

- Artur Sepp. Fourier transform for option pricing under affine jump-diffusions: An overview. 2003. URL www.hot.ee/seppar.
- Steven E. Shreve. *Stochastic Calculus for Finance II: Continuous-Time Models*. Springer Finance, 2004.
- Michael Stutzer. A simple nonparametric approach to derivative security valuation. *The Journal of Finance*, 51(5):1633–1652, 1996. ISSN 00221082. URL <http://www.jstor.org/stable/2329532>.
- J.E. Zhang and Jinghong Shu. Pricing s&p 500 index options with heston’s model. In *Computational Intelligence for Financial Engineering, 2003. Proceedings. 2003 IEEE International Conference on*, pages 85 – 92, 2003.

University of Cape Town