Two Approaches to Modelling the Volatility Skew

Chipo Masawi

Thesis presented in partial fulfilment of the requirements
for the degree of Master of Financial Mathematics

University of Cape Town
2008

Declaration

I, the undersigned, hereby declare that the work in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature: Signed by candidate

Date: 01/11/2008

Acknowledgements

I would like to take this opportunity to express my deepest gratitude to the following people:

My wonderful supervisor, Petrus Bosman from Cadiz. Thank you for your attention, guidance, insight and support during this research.

My mother and father , Shaleth and Francis Masawi, for always being an inspiration throughout my life. You have always supported my dreams and aspirations. I'd like to thank you for teaching me that I can do ALL things through Christ who gives me strength.

My sister and brother, Fadzai and Francis Junior. Thank you for support and love and most importantly for reminding me not to lose my sense of humour during this process.

My amazing friends namely, Kofi, Star, Spiwe, Rachel, Cynthia, Lorraine,Taph. I am eternally grateful for your generous support during this time.

I cannot end without acknowledging my late uncle David Aaron Mutsvanga. His unflinching courage and conviction will always inspire me. It is to him that I dedicate this work.

Abstract

In the last several years the volatility skew has come to play a vital role in option pricing, hedging and risk management. As a result, its modelling has become important. This study examines two approaches to modelling the volatility skew that is used to price options on the Johannesburg Stock Exchange (JSE) TOP40 index. The first approach involves using historical prices of the underlying index to obtain a model of the skew. Two models that use this approach, namely the Edgeworth and Normal Mixture AGARCH models were implemented. Although both models produced skew shapes that closely resemble those observed in South African equity markets, only the Edgeworth model can actually be used to price options. This is because, of the two models, it is the one that is able to back-out a risk-neutral density for the underlying. The second approach involves using the current prices of liquid options to model the skew. Two models, namely local volatility mixture and double exponential jump diffusion models were calibrated to the market skew. Of the two, the double exponential jump diffusion model was observed to produce the best fit to market implied volatilities.

Table of contents

1 Introduction

1.1 Subject of the thesis

Two approaches to modelling the volatility skew

1.2 Problem definition

The Black-Scholes option pricing model [45], pioneered in 1973 is used to price European options. It offers a relatively simple formula to compute the "fair" value of an option on an underlying asset given a limited number of inputs. The Black-Scholes model assumes a one-to-one relationship between the price of an option and its implied volatility i.e. the volatility fed into the model to yield the market price of an option. This means that the prices of options are often expressed in terms of their implied volatilities. The model also implies that this volatility remains constant throughout an option's lifetime. Therefore, options on the same underlying should also have the same Black-Scholes implied volatilities for all strikes and option maturities.

However, since the market crash of October 1987 [6], different Black-Scholes implied volatilities were observed for different strike prices and option maturities. The variation of implied volatilities with strike prices was termed a volatility smile or skew while variation with maturity is known as a term structure. Combining volatility smiles/skews and term structures gives rise to volatility surfaces. These surfaces provide a means for pricing an option with any strike price for any maturity. The general form of a volatility surface is illustrated in Figure 1.1.

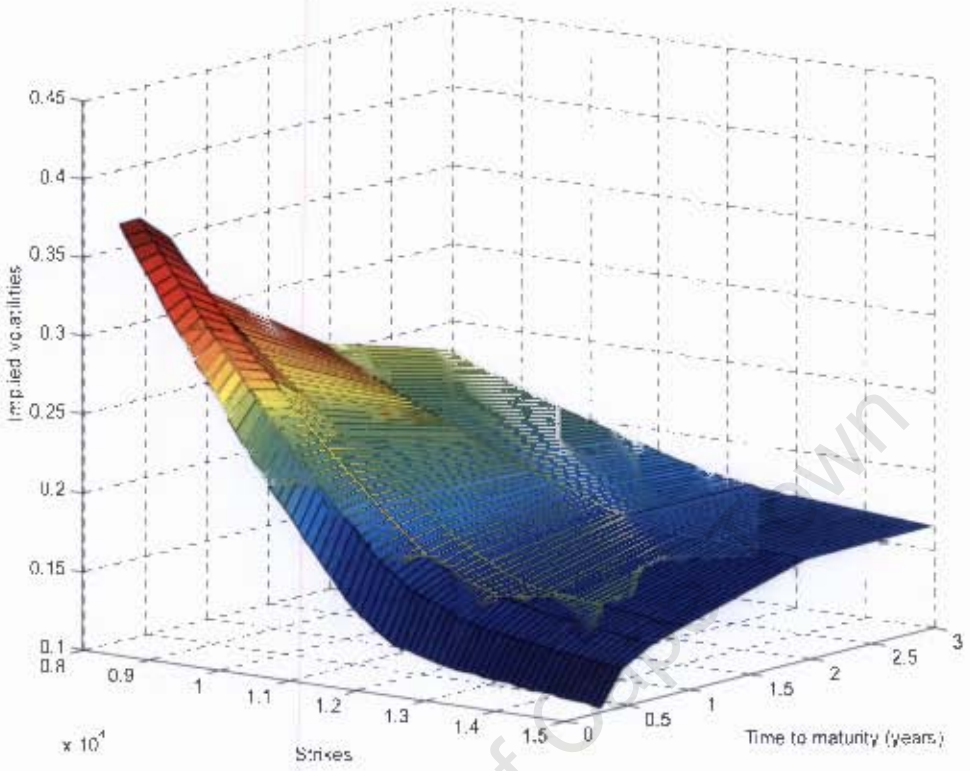


Figure 1.1: An illustration of a volatility surface

In some markets such as the foreign exchange market, the shape is parabolic. This results from deep-out-of-the-money and deep-in-the-money options having higher implied volatilities than at-the-money options [6]. The entire focus of this dissertation however is on the smile in equity index markets. In these markets the smile tends to be more negatively skewed and is therefore known as the volatility skew. The skew results from low strike price options i.e. out-of-the-money (OTM) puts and in-the-money (ITM) calls having higher implied volatilities than high-strike options i.e. ITM puts and OTM calls [6].

The volatility skew emerged because the assumptions upon which the Black-Scholes formula was based could not be justified empirically. An option's underlying asset returns are not normally distributed and their volatility is not constant as assumed by the formula [24]. In equity markets, the risk neutral

distribution of the underlying asset returns from which we can infer the probability of extreme events.

The latter models, on the other hand, are to be calibrated to market implied volatilities of options on the JSE TOPI40 index. The fitting quality of each model will then be evaluated and compared. These models would be particularly useful to banks or corporations. Once again, the probability distribution of the underlying index returns will be investigated.

1.4 Plan of development

Chapter 2 of this dissertation begins with a description of the models that make use of historical prices of the underlying asset to determine the volatility skew. These models namely the model that utilizes an EDGEWORTH expansion and the normal mixture GARCH are discussed in Sections 2.1 and 2.2 respectively. Chapter 3 contains a description of the normal mixture local volatility and double exponential jump diffusion models which are calibrated to the market implied volatility skew. The data and methodology used in this thesis are then discussed in Chapter 4. Model simulation and calibration results are presented in Chapter 5. Conclusions are then drawn on the basis of these findings in Chapter 6.

The CD-ROM attached this thesis includes all the programming code of the implementation of the four models reviewed. Modelling was implemented in Excel VBA.

2 Models calibrated to historical prices

2.1 Non-normal distributions and volatility modelling

Option pricing makes use of the risk-neutral valuation principle [1]. This principal states that the present value of an option is equal to the expected value of its payoff at expiry discounted to the present date at the risk-free rate [6]. The expectation is calculated using risk-neutral probabilities [1].

The Black-Scholes option pricing formula assumes that the underlying asset log returns are normally distributed. The skewness and kurtosis of a normal distribution is zero and three respectively [1]. Empirical evidence, such as that in [3], [4] and [5], however, suggests that the distribution of the log returns of financial time series:

- Are skewed (i.e. have non-zero skewness)
- Have fat tails (i.e. have a kurtosis greater than three)

This departure from normality led researchers such as Jarrow and Rudd [2] and Rubinstein [1] amongst others, to investigate alternative methods for valuing options. They investigated option pricing models where the distribution of the underlying asset log returns had different higher moments (i.e. skewness and kurtosis) from those of the normal distribution used in the Black-Scholes model.

An option pricing model that allows the higher moments (i.e. skewness and kurtosis) of the underlying asset log returns to differ from strict normality is implemented in this thesis. This model is referred to as the "Edgeworth" model in this thesis. It is now discussed.

2.1.1 The Edgeworth model

Jarrow and Rudd (1982) [2] derived an option pricing formula which made use of an Edgeworth expansion. They used this expansion because it allowed for additional flexibility over a distribution when pricing options [2]. This is because the expansion introduced skewness and kurtosis values of the empirical distribution (i.e. the returns density) as parameters into the option pricing model.

Rubinstein [1], on the other hand, simplified their approach by applying an Edgeworth expansion directly to discretised risk-neutral probabilities rather than the option pricing formula [2]. He then showed that the shape of the implied volatility smile is dependent on the assumed skewness and kurtosis of the distribution of the underlying asset returns [1].

probability distribution of the underlying asset returns has fatter left tails than a normal distribution. This implies that the market expects large price falls or market crashes to occur with a greater probability than is assumed by the Black-Scholes model i.e. the market expects that the probability of an extreme event such as a market crash is higher than the probability of the same event under a normal distribution.

Since asset managers or managers of an investment portfolio are concerned with pricing and hedging with options, option pricing models become an invaluable tool for computing the "fair" price of any option that they might wish to buy or sell. Now, asset managers who want to use the Black-Scholes model in a way which is consistent with the way the market place is using the model, would allow the volatility input to depend on its strike price (and time to maturity) [46]. They would therefore be interested in looking at the distribution of the implied volatilities across different strike prices i.e. the volatility skew. A theoretical model of the skew would provide a useful means of evaluating and comparing the prices of various options prior to calling a broker or going to the market. In particular, the implied volatilities could be compared to historic volatilities to determine how cheap or expensive an option is. A model of the volatility skew

2.1.1.1 Generating an Edgeworth density function

A standardised binomial distribution is first generated. Since this distribution is a discrete form of the normal distribution, it's mean must be equal 0, variance equal 1, skewness of 0 and kurtosis of 3. If n+1 are the number of equally spaced points required, at each point j = 0…..n, there is a random variable which is defined at these points and takes on values, $x(j)$ such that:

$$x(j) = \frac{2j - n}{\sqrt{n}}.$$

The random variable has an associated probability $b(x(j))$ which is equal to:

$$b(x(j)) = \frac{n!}{j!(n-j)!}(1/2)^n.$$

The standardised binomial probabilities $b(x(j))$ are then transformed via Edgeworth expansion terms to form the adjusted, standardised probability distribution, $f(x(j))$ as follows:

$$f(x(j)) = c(x(j)) * b(x(j)).$$

$f(x(j))$ is also referred to as the standardised "Edgeworth density" [1] and the term c is known as an Edgeworth factor or expansion. This factor is a polynomial that accounts for the effects of departure from normality. It is computed as follows:

$$c(x) = 1 + (1/6)\xi(x^3 - 3x) + (1/24)(\kappa - 3)(x^4 - 6x^2 + 3)$$
$$+ (1/72)\xi^2(x^6 - 15x^4 + 45x^2 - 15).$$

ξ and κ are the desired skewness and kurtosis respectively. The expansion is only an approximation and as a result, $\sum_j f(x(j)) \neq 1$.Additionally, there are slight

errors in the moments [1]. The distribution is then rescaled by replacing $f(x(j))$ with $f'(x(j))$ such that $f'(x(j)) = \frac{f(x(j))}{\sum_j f(x(j))}$. The mean and variance of this new distribution are then calculated as:

$$Mean = \sum_j f'(x(j))x(j).$$
$$Variance = \sum_j f'(x(j))(x(j) - Mean)^2.$$

Once the mean and variance of the $f'(x(j))$ distribution are calculated, the random variables $x(j)$ are transformed to Edgeworth variables, $x'(j)$ as follows:

$$x'(j) = \frac{x(j) - Mean}{\sqrt{Variance}}.$$

This transformation ensures that the Edgeworth distribution has mean 0 and variance 1. If a large value of n is chosen (i.e. n greater than 100) the binomial distribution will approximate a normal distribution more closely and the skewness and kurtosis of the distribution should be very similar to those initially chosen.

Problems that arise from this distribution include obtaining values for $f'(x(j))$ that are negative. Since $f'(x(j))$ is a density function, it cannot have negative values. As a result, it becomes essential to define a range or locus of (ξ, κ) values for which $f'(xj) > 0$. Another problem that could occur is that even is $f'(xj) > 0$, the density may not be unimodal. Once again it becomes necessary to define (ξ, κ) values for which the distribution is unimodal.

When the Edgeworth expansion fails to produce positive unimodal distributions, the locus of (ξ, κ) pairs can be enlarged by using the Gram-Charlier expansion [1]. Although this expansion is very similar to the Edgeworth expansion, for computational process it performs better than the Edgeworth expansion [48]. This is because the Gram-Charlier omits the last term of the Edgeworth expansion and

thus simplifies the search for the domain where the distribution is positive [48]. The Gram-Charlier expansion has also been used in option pricing by other researchers such as Jarrow and Rudd (1982) [2], Madan and Milne (1994) [7] and Corrado and Su (1997) [8]. It now expresses the Edgeworth factor, c, as follows:

$$c(x) = 1 + (1/6)\xi(x^3 - 3x) + (1/24)(\kappa - 3)(x^4 - 6x^2 + 3).$$

The locus of (ξ, κ) pairs can also be influenced by the number of discrete points, n required to create the Edgeworth distribution [1]. The greater the number of points, the smaller the locus will be. (ξ, κ) pairs which allow for positive unimodal distributions (for n=100) are shown in table A1 in appendix A.

The second step in Rubinstein's option valuation method involves using the Edgeworth distribution to calculate option prices. This process is now discussed.

2.1.1.2 Option valuation

In order to derive an option valuation formula, the standardised distribution, must be adjusted to ensure that it has the appropriate risk neutral mean and variance [3]. The first step in calculating option prices is to derive a formula for the risk-neutral mean of the logarithm of the underlying asset price. Let:

S_t = current underlying asset price

r = annualised risk free rate

t = time to expiry

μ = annualised risk-neutral mean of log returns

σ = annualised risk-neutral volatility of log returns

$P_j = f'(xj)$ = risk-neutral probabilities associated with asset price Sj at expiry

q =continuous dividend yield

Given r, q, t and σ, the risk neutral drift, μ, can be calculated as follows [1]:

The possible values of the asset in this model at expiry are denoted by S_j. Where S_j can be calculated from the Edgeworth variables, $x'(j)$ as follows [1]:

$$S_j = S_t \exp\left(\mu t + \sigma\sqrt{t}x'(j)\right) \quad \textbf{(2.1)}$$

To value an option on an asset that pays out a continuous dividend yield q, the expected growth rate of the asset is set equal to r-q in the risk neutral measure. This means that the expected asset price at expiry is $S_t e^{(r-q)t}$. Consequently, if $\sum_j P_j = \sum_j f'(xj) = 1$, then:

$$e^{(r-q)t} = \sum_j P_j (S_j / S_o)..$$

It then follows that:

$$e^{(r-q)t} = \sum_j P_j \exp\left(\mu t + \sigma\sqrt{t}x'(j)\right) = \left(\sum_j P_j \exp\left(\sigma\sqrt{t}x'(j)\right)\right)\exp(\mu t)$$

$$t(r-q) = \log\left(\sum_j P_j \exp\left(\sigma\sqrt{t}x'(j)\right)\right) + \mu t$$

$$\mu = (r-q) - \log\left(\sum_j P_j \exp(\sigma\sqrt{t}x'(j)\right)/t$$

Once the risk neutral drift is computed, option prices are then calculated by discounting the expected value of their payoffs at expiry using the risk-free rate. The value of call options with strikes, K_i , for $i = 1,.....,m$ on an underlying asset S_t , is given as follows:

$$\text{Call price} = \exp(-r*t)*\sum_j P_j \max(0, S_j - K_i) \quad \textbf{(2.2)}$$

Put-Call parity can be used to compute the corresponding put values. Implied volatilities can be calculated from these call prices using the Black-Scholes formula. By repeatedly calculating call values and implied volatilities for a range of strike prices, volatility skew shapes can be generated for different values of skewness and kurtosis. Rubinstein [1] observed that symmetric risk-neutral Edgeworth densities produced symmetric smile patterns around ATM levels. Skewed distributions on the other hand gave rise to skewed or asymmetric smiles. Three implied volatility skew shapes produced for different skewness and kurtosis values are illustrated in Figure 2.1.

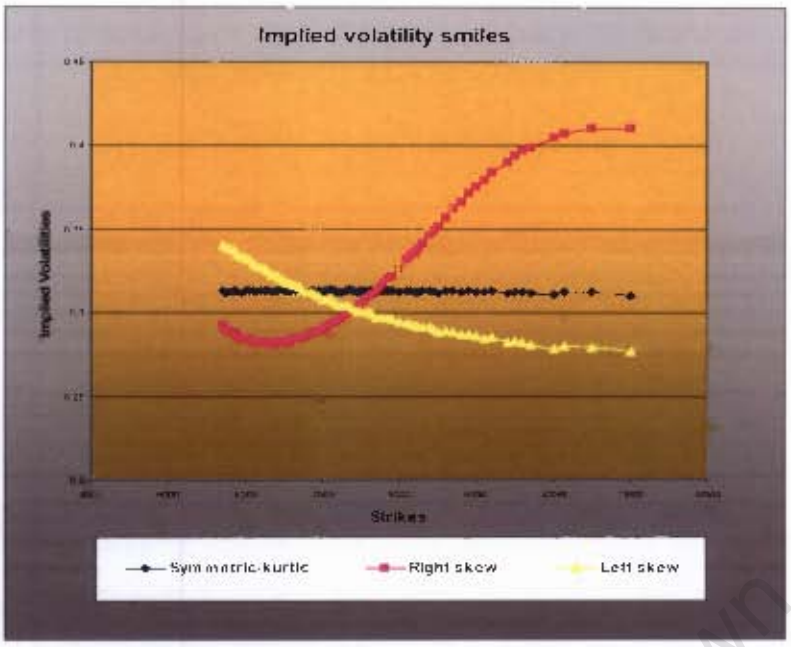


Figure 2.1: Implied volatility shapes for the following parameter values:
Symmetric-kurtic $\xi = 0, \kappa = 3$, Right skew; $\xi = 0.6, \kappa = 4.8$, Left skew: $\xi = -0.39, \kappa = 3.4$

Figure 2.1 illustrates the different implied volatility shapes obtained for different skewness and kurtosis values.

2.1.2 Advantages and disadvantage of the model

The Edgeworth model makes use of the skewness and kurtosis of the underlying asset. It can also produce volatility skew shapes that closely resemble those observed in equity markets. Furthermore, risk neutral densities of the underlying log returns can be backed out from this model. These densities can be compared with those observed in the market to see if there is any relative value to trading using the Edgeworth model over time.

The Edgeworth model, however, assumes a constant volatility for the log returns. The volatility exhibited in financial data, however, varies with time. In order to provide a more realistic model for the volatility dynamics of the underlying asset returns, Generalised Autoregressive Conditional Heteroskedasticity (GARCH) models were investigated. These models which model volatility directly and therefore allow for time-varying volatilities are now discussed.

2.2 GARCH option pricing models

Generalised autoregressive conditional heteroskedasticity or GARCH models, for short were pioneered by Bollerslev [36] in 1986. They have been used successfully in modelling the time varying volatility exhibited in financial data. In order to provide a more realistic model for the volatility dynamics of the underlying asset, GARCH modeling has been incorporated into option pricing models.

Alternative models to Black-Scholes, in which the underlying asset is assumed to follow a GARCH (1, 1) process, have been proposed by Duan (1995, 1999) [37], [38] and Heston and Nandi (2000) [39].

Most recently, a class of models known as normal mixture GARCH (1, 1) models has been examined by several authors such as Haas et al. [42] and Alexander [41]. In these models, the distribution of the GARCH error processes is generated by a mixture of normal distributions.

Normal mixture distributions are probability weighted sums of normal distributions. For example, given two normal distributions, $\Phi_1(x) = \Phi(x; \mu_1, \sigma_1^2)$ and $\Phi_2(x) = \Phi(x; \mu_2, \sigma_2^2)$, a normal mixture density function, g(x), is:

$$g(x) = p\Phi_1(x) + (1-p)\Phi_2(x).$$

where μ_i and σ_i^2 represent the means and variances respectively of the individual normal distributions respectively and p is a positive constant. These parameters determine the shape and scale of the normal mixture distribution. This property is illustrated in Figures 2.2 and 2.3.

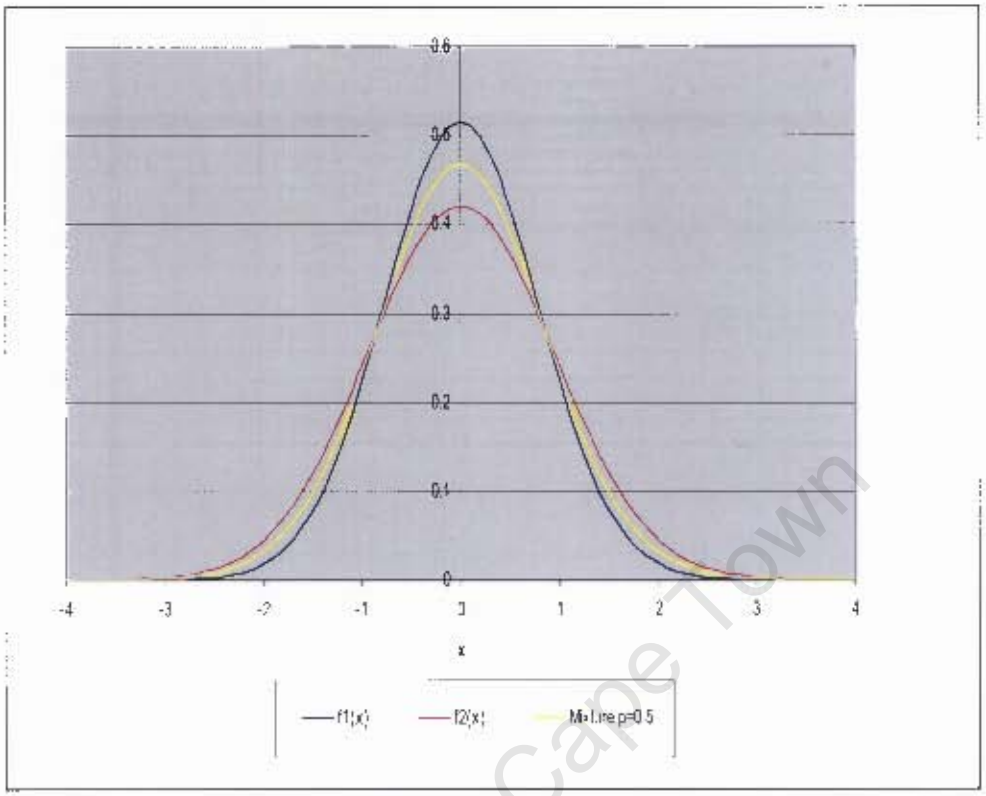


Figure 2.2: Normal mixture density with zero mean

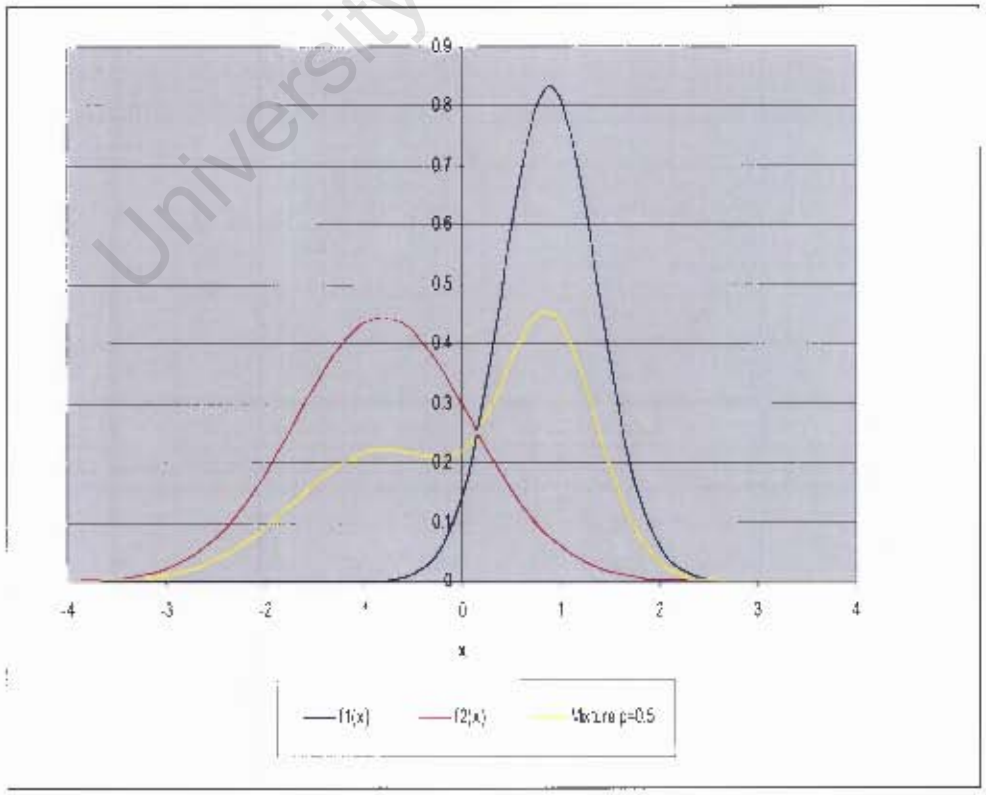


Figure 2. 2.3 : Normal mixture density with different means

A mixture with two normal distributions each with zero means will give rise to a distribution with zero skew that has fatter tails than the one distribution but not the other. On the other hand a mixture of two normal distributions with different means produces thin tailed, skewed distributions. Normal mixture distributions with zero means are therefore particularly useful in modeling financial returns data which often have fatter tails than normal distributions because they too are highly leptokurtic [24].

Normal mixture GARCH (1,1) models give rise to skewed leptokurtic distributions. They are also able to describe the time variation in the conditional skewness and kurtosis of financial returns data [41]. The implied volatility skew, which arises partly from leptokurtosis in the distribution of the underlying asset returns, is also known to exhibit variation over time. As a result, normal mixture GARCH (1, 1) models are consistent with observed skew dynamics [41].

Recently, Alexander and Lazar introduced asymmetry, in the form of GJR (Glosten, Jagannathan and Runkle) and AGARCH (asymmetric GARCH) component variance processes, into the normal mixture GARCH (1, 1) model [40]. This asymmetry introduces an additional source of skewness (known as the leverage effect, see [24]) in the returns distribution which enables the model to be successfully applied to equity indices. On examining the model's ability to simulate the volatility skew of options on an equity index, they concluded that asymmetric normal mixture GARCH (1, 1) models produced the most realistic equity skews above symmetric normal mixture GARCH (1,1) models and GARCH models with single variance components.

The asymmetric normal mixture GARCH (1, 1) model is studied and used for option pricing in this thesis. The model is discussed in Section 2.2.1.

2.2.1 The asymmetric normal mixture GARCH (1, 1) model

The normal mixture GARCH (1, 1) model with K mixtures is denoted by NM (K)-GARCH (1, 1). $y = \varepsilon_t$ is assumed to be the conditional mean equation in this model. The error term, ε_t in this equation is also assumed to have a normal mixture density i.e. [40]:

$$\varepsilon_t \mid \mathrm{I}_{t-1} \square\ NM(p_1,....,p_K,\mu_1,....,\mu_K,\sigma_{1t}^2,....,\sigma_{Kt}^2), \quad \sum_{i=1}^{K} p_i = 1\ ,\ \sum_{i=1}^{K} p_i\mu_i = 1\ .$$

That is, the density of the error term can be written as a probability p (where $p_i > 0$) weighted sum of K (≥ 2) normal density functions each with different constant means $(\mu_1,....,\mu_K)$ and time-varying variances $(\sigma_{1t}^2,....,\sigma_{Kt}^2)$ .The density function of the error term can be expressed as:

$$\eta(\varepsilon_t) = \sum_{i=1}^{K} p_i\varphi_i(\varepsilon_t)\ .$$

Where φ_i represent normal density functions. In [41], the variance of each normal density is assumed to follow either the AGARCH (asymmetric GARCH) process or the GJR process. These processes enable asymmetry to be incorporated into the normal mixture density. One possibility is considered, the NM (K)-AGARCH, where the variance of each mixture is expressed as:

$$\sigma_{it}^2 = \omega_i + \alpha_i(\varepsilon_{t-1} - \lambda_i)^2 + \beta_i\sigma_{it-1}^2 \quad \text{for } i = 1,....,K\ . \qquad \textbf{(2.3)}$$

λ represents the leverage effect i.e., the correlation between asset returns and volatility. a , β *and* ω are the usual GARCH constants, see [36].

The overall conditional variance is then expressed as [40]:

$$\sigma_t^2 = \sum_{i=1}^{K} p_i \sigma_{it}^2 + \sum_{i=1}^{K} p_i \mu_i^2 .$$

In order to produce non-negative variances, the following set of conditions must be obeyed for $i = 1,.....K$:

$$0 < p_i < 1 , \quad i = 1,.....K-1, \quad \sum_{i=1}^{K-1} p_i < 1, \quad 0<\alpha_i , \quad 0 \le \beta_i \le 1$$

Additionally, all individual long term variances must be finite and positive. If expectations are taken on both sides of equation 2.3, the following relationship between individual and long term variances, σ_{it}^2 and σ_t^2, respectively, is obtained:

$$E(\sigma_{it}^2) = \omega_i + \alpha_i(\varepsilon_{t-1} - \lambda_i)^2 + \beta_i \sigma_{it-1}^2$$

$$E(\sigma_{it}^2) = \omega_i + \alpha_i E(\varepsilon_t^2) + \alpha_i \lambda_i^2 + \beta_i E(\sigma_{it}^2) .$$

$$E(\sigma_{it}^2)(1-\beta_i) = (\omega_i + \alpha_i \lambda_i^2) + \alpha_i E(\varepsilon_t^2)$$

For an NM (K)-GARCH (1, 1), the following equation represents the overall long term variance [41]:

$$E(\varepsilon_t^2) = E(\sigma_t^2) = \frac{\sum_{i=1}^{K} p_i \mu_i^2 + \sum_{i=1}^{K} \frac{p_i(\omega_i)}{1-\beta_i}}{\sum_{i=1}^{K} \frac{p_i(1-\alpha_i-\beta_i)}{1-\beta_i}} \quad \textbf{.(2.4)}$$

The overall long term variance for the asymmetric NM (K) –AGARCH (1, 1) can therefore be expressed as:

$$E(\varepsilon_t^2) = E(\sigma_t^2) = \frac{\sum_{i=1}^{K} p_i \mu_i^2 + \sum_{i=1}^{K} \frac{p_i(\omega_i + \alpha_i \lambda_i^2)}{1-\beta_i}}{\sum_{i=1}^{K} \frac{p_i(1-\alpha_i-\beta_i)}{1-\beta_i}} \quad . \textbf{(2.5)}$$

Parameters in the NM (K) –AGARCH (1, 1) model must be chosen so that non-negative overall variances are obtained. From equation 2.5, this corresponds to the following conditions being satisfied:

$$m = \sum_{i=1}^{K} p_i \mu_i^2 + \sum_{i=1}^{K} \frac{p_i(\omega_i + \alpha_i \lambda_i^2)}{1-\beta_i} > 0$$

$$n = \sum_{i=1}^{K} \frac{p_i(1-\alpha_i-\beta_i)}{1-\beta_i} > 0 \quad .$$

Additionally, positive individual variances need to be obtained. Therefore a further condition must be satisfied, i.e.:

$$E(\sigma_{it}^2) = \frac{(\omega_i + \alpha_i \lambda_i^2) + \alpha_i E(\varepsilon_t^2)}{(1-\beta_i)} > 0$$

$$\therefore \qquad \omega_i + \alpha(_i \lambda_i^2 + E(\varepsilon_t^2)) > 0$$

From conditions *m* and *n* above , the following is obtained:

$$r = \omega_i + \alpha_i(\frac{m}{n} + \lambda_i^2) > 0$$

The procedure used to estimate GARCH model parameters is now discussed.

2.2.1.1 Parameter Estimation

GARCH model parameters are estimated by maximum likelihood methods from historical data. The idea behind this procedure is to obtain GARCH parameter values that maximize the likelihood or probability of the data occurring under an assumption about its distribution.

The likelihood of an observation x on a random variable is defined as the value of its probability density function, given as $f(x,\theta)$. $\theta = (\theta_1, ...\theta_m)$ represents the set of parameters of the density function [24]. Supposing that m observations are made on the same random, the likelihood function, $L(\theta|x_1, x_2....x_n)$ of the set of observations $(x_1, x_2, ...x_m)$ is the product of the likelihoods at each point that the observation is made. I.e. [24]:

$$L(\theta|x_1, x_2....x_m) = \prod_{i=1}^{m} f(x_i, \theta) \quad . \quad \textbf{(2.6)}$$

Using maximum likelihood methods, the best estimate of $\theta = (\theta_1, ...\theta_m)$ is the value that maximizes the likelihood function in equation 2.6. Maximising the likelihood also equates to maximising its logarithm. Therefore the MLE is usually found to be the value of $\theta = (\theta_1, ...\theta_m)$ that maximises the log-likelihood. Taking logarithms of the expression in equation 2.6 gives rise to the log-likelihood function:

$$\ln L(\theta|x_1, x_2....x_m) = \sum_{i=1}^{m} \ln f(x_i, \theta) \quad . \quad \textbf{(2.7)}$$

Once GARCH model parameters are estimated via this likelihood function, the model can be used for option pricing. The option pricing process is discussed in the following section.

2.2.1.2 Option pricing

The price of a European option is its risk neutral expected payoff at maturity discounted to the present date. This expectation is taken over the probability density of the underlying asset. If the density function of the underlying asset is a mixture of distributions, the price of the option on the underlying asset becomes a weighted sum of the prices under each density function in the mixture [24].

To price an option using a single mixture normal GARCH (1,1) model, the Black-Scholes formula is used with volatility input given as the square root of the expected variance rate over the life of the option. Suppose today is day n and the option expires on day n+N. The expected variance rate during the life of the option is given as [4]:

$$\frac{1}{N}\sum_{k=0}^{N-1} E\left[\sigma_{n+k}^2\right].$$

The longer the life of the option, the closer this value is to the long-term variance, V, given as $\frac{\omega}{1-\alpha-\beta-\alpha\lambda^2}$ for an AGARCH process.

Normal mixture GARCH (1,1) European option prices are therefore calculated as weighted sums of these Black-Scholes prices [41].

2.2.2 Limitations of the models

Asymmetric normal mixture GARCH (1,1) models produce skews that closely resemble those observed in the equity markets [24] and also give rise to skewed leptokurtic densities. These densities, however, are not risk neutral and therefore the skew produced by these models cannot actually be used to price options using the Black Scholes formula.

The asymmetric normal mixture GARCH (1,1) and Edgeworth models offer an alternative method of modelling the volatility skew. In these models, the skew is estimated from the historical prices of the underlying asset. Current prices of liquid options can also be used to determine the volatility skew. Models that are calibrated to market implied volatilities are now discussed.

3 Models calibrated to market implied volatilities

3.1 Deterministic local volatility models

In the Black-Scholes framework the underlying asset, S_t is assumed to evolve log normally according to the diffusion process:

$$dS_t = \mu S_t dt + \sigma S_t dW \quad . \textbf{(3.1)}$$

where μ is defined as the expected continuously compounded rate of return, σ the underlying asset volatility and dW is a standard Brownian motion with mean zero and variance dt. Both μ and σ parameters are assumed to be constant. The diffusion process in equation 3.1 is used in the derivation of the Black-Scholes option pricing formula.

Implied volatilities are estimates of volatility from the market. When these volatilities are calculated from traded options, they are found to be non-constant. This contradicts the assumption of constant volatility in the Black-Scholes model. In fact, implied volatilities are observed to be dependent on both the strike price and time to maturity of the option. This gives rise to a volatility smile or skew as well as a term structure [18].

Several approaches have been suggested for dealing with this phenomenon. One approach was to allow the volatility in the Black-Scholes diffusion process i.e. equation 3.1, to be time dependent. This was done by Merton [15] in 1973.In his model, the asset price process is governed by the following stochastic differential equation [18]:

$$dS_t = \mu S_t dt + \sigma(t) S_t dW \quad . \textbf{(3.2)}$$

Although this model accounts for the time dependence of implied volatilities, it does not account for their dependence on the strike price. Other approaches that were tried included the jump diffusion model of Merton [17] and modelling volatilities as correlated stochastic processes as in the models of Hull and White

[16] and Heston [12], [30]. These models were able to explain the volatility skew but their drawback was that they introduced non-traded sources of risk. As a result, the models do not allow for arbitrage free pricing or hedging.

This observation led to the introduction of a class of models known as local volatility models. These models present an alternative way of fitting the implied volatility skew. Their advantage over stochastic and jump diffusion models is that apart from matching the skew, they also offer risk neutral pricing like the Black-Scholes model.

In this new framework, the asset price dynamics in equation 3.1 are now expressed as the following non-linear diffusion process:

$$dS_t = \mu S_t dt + \sigma(S_t, t) S_t dW_t \quad \textbf{(3.3)}$$

The parameter $\sigma(S,t)$ is known as the local volatility. The local volatility parameter in these models is a deterministic function of both the asset price and time [11], [10], and [14]. Local volatility models were first introduced by Dupire [11], Derman and Kani [10] and Rubinstein [14]. They derived binomial and trinomial tree based algorithms where local volatilities were backed out through forward induction from the implied volatilities of market traded vanilla options [25],[26].

By applying an approach similar to Dupire [11], [18] another class of local volatility models was introduced by Brigo and Mercurio [20]. These local volatility models are known as mixture models. They are discussed in Section 3.1.1.

3.1.1 Local volatility mixture models

Brigo and Mercurio [19], [20] proposed local volatility mixture models as an alternative to the Black-Scholes model. Their models are based on the assumption that the underlying asset price density is a mixture of known basic densities. Not only do these models fit market data well but they also result in analytical formulas for European options and are flexible enough to recover a large variety of implied volatility shapes [19].

In their mixture models, it is assumed that a risk-neutral measure exists. The dynamics of the underlying asset, S, under this measure, namely the forward measure Q^T, can be expressed as follows [21]:

$$dS_t = \mu S_t dt + \sigma(S_t, t) S_t dW_t \quad . \quad \textbf{(3.4)}$$

where μ is a constant and W is a Q^T standard Brownian motion. $\sigma(S_t, t)$ is the local volatility function. In order to ensure the existence of a strong solution to equation 3.4, the local volatility function must satisfy the following linear-growth condition [21]:

$$\sigma^2(y, t) y^2 = L(1 + y^2) \quad . \quad \text{uniformly in t}$$

where L is a suitable positive constant. Also considered in the local volatility mixture models are N diffusion processes with the following dynamics [21]:

$$dS_t^i = \mu S_t^i dt + v_i(S_t^i, t) dW_t \quad . \quad \textbf{(3.5)}$$

with the same μ for each i where $S_0^i = S_0$ and $v_i(S_t^i, t)$ represent real functions that are chosen to ensure the uniqueness and existence of a solution to equation 3.5 [21].

Once again, it is assumed that for i= 1,…., N and for a constants L_i, the following linear growth condition holds:

$$v_i^2(y,t) = L_i(1+y^2) \quad . \quad \text{uniformly in t}$$

Now let $p_t^i(\cdot)$ be the density functions of S_t^i where each density function obeys the relationship:

$$p_t^i(y) = \frac{d}{dy} Q^T \{S_t^i \leq y\}.$$

Brigo and Mercurio [13] were able to derive the local volatility function $\sigma(S_t,t)$ in equation 3.4 in such a way that for each t, the Q^T density function of the underlying asset could be expressed as a combination of these basic densities i.e.:

$$p_t(y) = \frac{d}{dy} Q^T \{S_t \leq y\}$$

$$= \sum_{i=1}^{N} \lambda_i \frac{d}{dy} Q^T \{S_t^i \leq y\}$$

$$= \sum_{i=1}^{N} \lambda_i p_t^i(y)$$

i.e. $$p_t(y) = \sum_{i=1}^{N} \lambda_i p_t^i(y) \quad . \quad \textbf{(3.6)}$$

Where the $\lambda_i > 0$ and $\sum_i^N \lambda_i = 1$. It can be shown that $p_t(\cdot)$ is a proper Q^T density function as follows:

By definition:

$$\int_0^\infty y p_t(y)dy = \sum_{i=1}^N \lambda_i \int_0^\infty y p_t^i(y)dy$$

To solve $\int_0^\infty y p_t^i(y)dy$, the diffusion process in equation 3.5 must be solved

Consider the diffusion process:

$$dS_t^i = \mu S_t^i dt + v_i(S_t^i, t)S_t^i dW_t$$

Define $Y_t^i = \ln S_t^i$

$$dY_t^i = \frac{1}{S_t^i} dS_t^i - \frac{1}{2(S_t^i)^2}(dS_t^i)^2$$

$$= [\mu - \frac{1}{2}v_i^2(S_t^i, t)]dt + v_i(S_t^i, t)dW_t$$

And thus

$$Y_t^i = Y_0^i + \int_0^t [\mu - \frac{1}{2}v_i^2(S_t^i, u)]du + \int_0^t v_i(S_t^i, u)dW_u$$

The equation for S_t^i in equation 3.5 is therefore:

$$S_t^i = S_o^i \exp\left\{\int_0^t [\mu - \frac{1}{2}v_t^2\left(S_t^i, u\right)]du + \int_0^t [v_t\left(S_t^i, u\right)]dW_u\right\}$$

$$= S_o^i \exp\left\{\int_0^t \mu ds + \int_0^t v_t\left(S_t^i, s\right)]dW_s\right\}.$$

Thus if v_t is deterministic, S_t^i is lognormally distributed with mean $S_o^i e^{\mu t}$. From this result,

$$\int_0^\infty y p_t^i(y)dy = S_0^i e^{\mu t}$$

$$\therefore \int_0^\infty y p_t(y)dy = \sum_{i=1}^N \lambda_i \int_0^\infty y p_t^i(y)dy$$

$$= \sum_{i=1}^N \lambda_i S_0^i e^{\mu t}$$

$$= \sum_{i=1}^N \lambda_i S_0 e^{\mu t}$$

$$= S_0 e^{\mu t}$$

Since $S_0^i = S_0$ and $\sum_{i=1}^N \lambda_i = 1$

As in Dupire's work [12], [18], the σ coefficient in equation 3.4 is extracted by applying the Fokker-Planck Equation [43] to S with $p_t(\cdot)$ given in equation 3.6, such that:

$$\frac{\partial}{\partial t} p_t(y) = -\frac{\partial}{\partial y}(\mu y p_t(y)) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(\sigma^2(y,t)y^2 p_t(y))$$

Each density $p_t^i(y)$ satisfies the following Fokker Planck equation:

$$\frac{\partial}{\partial t} p_t^i(y) = -\frac{\partial}{\partial y}(\mu y p_t^i(y)) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(v_i^2(y,t) p_t^i(y))$$

From these equations, an expression for $\sigma(y,t)$ which is consistent with the marginal density in equation 3.6 can be derived as follows:

If $\frac{\partial}{\partial t} p_t(y) = -\frac{\partial}{\partial y}(\mu y p_t(y)) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(\sigma^2(y,t) y^2 p_t(y))$

and $p_t(y) = \sum_{i=1}^{N} \lambda_i p_t^i(y)$

Then $\frac{\partial}{\partial t}\sum_{i=1}^{N} \lambda_i p_t^i(y) = -\frac{\partial}{\partial y}(\mu y \sum_{i=1}^{N} \lambda_i p_t^i(y)) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(\sigma^2(y,t) y^2 \sum_{i=1}^{N} \lambda_i p_t^i(y))$ (*)

But $\frac{\partial}{\partial t} p_t^i(y) = -\frac{\partial}{\partial y}(\mu y p_t^i(y)) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(v_i^2(y,t) p_t^i(y))$

$\therefore \quad \frac{\partial}{\partial t}\sum_{i=1}^{N} \lambda_i p_t^i(y) = -\frac{\partial}{\partial y}(\mu y \sum_{i=1}^{N} \lambda_i p_t^i(y)) + \frac{1}{2}\frac{\partial^2}{\partial y^2}(v_i^2(y,t) \sum_{i=1}^{N} \lambda_i p_t^i(y))$ (**)

But if expressions (*) and (**) are to be identical then

$$v_i^2(y,t)\sum_{i=1}^{N} \lambda_i p_t^i(y) = \sigma^2(y,t) y^2 \sum_{i=1}^{N} \lambda_i p_t^i(y)$$

i.e. $$\sigma^2(y,t) = \frac{v_i^2(y,t)\sum_{i=1}^{N} \lambda_i p_t^i(y)}{y^2 \sum_{i=1}^{N} \lambda_i p_t^i(y)}$$

This result implies that:

$$\sigma(y,t) = \sqrt{\frac{\sum_{i=1}^{N} \lambda_i v_i^2(y,t) p_t^i(y)}{\sum_{i=1}^{N} \lambda_i y^2 p_t^i(y)}} \quad . \qquad \textbf{(3.7)}$$

The new dynamics for S under the forward measure Q^T can therefore be expressed as follows:

$$dS_t = \mu S_t dt + \sqrt{\frac{\sum_{i=1}^{N} \lambda_i v_i^2(y,t) p_t^i(y)}{\sum_{i=1}^{N} \lambda_i y^2 p_t^i(y)}} S_t dW_t \quad \textbf{(3.8)}$$

The stochastic differential equation in equation 3.8 describes dynamics for the underlying asset price that lead to the marginal density in equation 3.6. It can be observed that these dynamics are now dependent on known basic densities $p_t^i(y)$ and their $v_i(t, S_t^i)$ diffusion coefficients. Now, assuming that the dynamics for each S_t^i in equation 3.5 leads to explicit European option prices and the dynamics in equation 3.8 have a unique strong solution, the time zero price of a European option with maturity T, strike K, written on the asset S, can be expressed as [19] :

$$EO(K,T,\omega) = P(0,T)E^T\left\{ \left[\omega(S_T - K)\right]^+ \right\}$$

$$= \sum_{i=1}^{N} \lambda_i EO_i(K,T,\omega) \qquad \textbf{(3.9)}$$

where $\omega = 1$ for a call, $\omega = -1$ for a put, E^T represents the expectation under Q^T and EO_i the European option price associated with the underlying S_t^i whose

dynamics are given in equation 3.5 [19]. $P(t,T)$ represents the time t price of a zero coupon bond maturing at time T.

Therefore, starting from analytically tractable basic densities $p_t^i(y)$, Brigo and Mercurio [19]-[23] were able to derive models that preserved this property. In fact the model discussed can be applied to any mixture of known distributions. Brigo and Mercurio [21] reviewed in great detail, the cases when the asset price density is given as:

- A mixture of lognormal densities with equal means
- A shifted mixture of lognormal distributions
- A mixture of lognormal densities which allows for different means in the individual distributions
- A mixture of hyperbolic-sine processes

The first three cases, which are implemented in this thesis, are now discussed briefly.

3.1.1.1 Log normal mixture model

In this model, the densities $p_t^i(\cdot)$ are all lognormal. Each density is assumed to have the same mean, μ. It is assumed that for each i:

$$v_i(y,t) = \sigma_i(t)y$$

Where σ_i's, the volatilities of the basic densities, are deterministic functions of t defined on the interval [0, T*], T* >0 on a given time horizon. They are also bounded from above and below by positive constants [20]. The diffusion process for each mixture component is therefore:

$$dS_t^i = \mu S_t^i dt + \sigma_i(t) S_t^i dW_t .$$

Marginal densities for each S_t^i conditional on S_0 are then given by:

$$p_t^i(y) = \frac{1}{yV_i(t)\sqrt{2\pi}} \exp\left\{ -\frac{1}{2V_i^2(t)} \left[\ln(\frac{y}{S_0}) - \mu t + \frac{1}{2} V_i^2(t) \right]^2 \right\}$$

Where the root mean squared volatility of the basic densities are given as [20]:

$$V_i(t) = \sqrt{\int_0^t \sigma_i^2(u)du}$$

Lognormal densities are chosen due to their analytical tractability and obvious link to the Black-Scholes model. Also mixtures of lognormals produce more leptokurtic distributions that Gaussian distributions [24]. Since log returns are leptokurtic, this is an ideal feature to incorporate into an option pricing model. Additionally, empirical work such as that done by Ritchey (1990), Melick and Thomas (1997), Bhupinder (1998), Guo (1998), Alexander and Narayanan (2001) [21], all suggest that lognormal mixtures work well to reproduce market volatility structures.

The following propositions are proved in [23].

Proposition 3.1.1 [19]:

It is assumed that the σ_i's are continuous and bounded from below by a positive constant and that there exists an $\varepsilon > 0$ such that $\sigma_i = \sigma_0 > 0$ for each t in $[0, \varepsilon]$ and for i= 1,...., N .Let:

$$v(y,t) = \sqrt{\frac{\sum_{i=1}^{N} \lambda_i \sigma_i^2(t) \frac{1}{V_i(t)} \exp\left\{-\frac{1}{2V_i^2(t)}\left[\ln(\frac{y}{S_0}) - \mu t + \frac{1}{2}V_i^2(t)\right]^2\right\}}{\sum_{i=1}^{N} \lambda_i \frac{1}{V_i(t)} \exp\left\{-\frac{1}{2V_i^2(t)}\left[\ln(\frac{y}{S_0}) - \mu t + \frac{1}{2}V_i^2(t)\right]^2\right\}}}$$

For (y,t)>(0,0) and $v(y,t) = \sigma_0$ for (y,t)= (S_0,t). Then the SDE:

$$dS_t = \mu S_t dt + v(S_t, t) S_t dW_t \quad \textbf{(3.10)}$$

has a unique strong solution. Its marginal density is:

$$p_t(y) = \sum_{i=1}^{N} \lambda_i \frac{1}{yV_i(t)\sqrt{2\pi}} \exp\left\{-\frac{1}{2V_i^2(t)}\left[\ln(\frac{y}{S_0}) - \mu t + \frac{1}{2}V_i^2(t)\right]^2\right\}$$

Additionally, for each t, y > 0

$$0 \leq \tilde{\sigma} \leq v(y,t) \leq \hat{\sigma} \leq +\infty$$

Where

$$\tilde{\sigma} := \inf_{t \geq 0} \{\min_{i=1,\ldots,N} \sigma_i(t)\}$$

$$\hat{\sigma} := \sup_{t \geq 0} \{\max_{i=1,\ldots,N} \sigma_i(t)\}$$

Proposition 3.1.2 [19]:

Consider a European option $EO(K,T,\omega)$, written on the asset with the dynamics (3.4), with maturity T and strike K. Its value at time zero is given as:

$$EO(K,T,\omega) =$$

$$\omega P(0,T) \sum_{i=1}^{N} \lambda_i \left[S_0 \exp(\mu T) \Phi\left(\omega \frac{\ln(\frac{S_0}{K}) + (\mu + \frac{1}{2}\eta_i^2)T}{\eta_i \sqrt{T}} \right) - K\Phi\left(\omega \frac{\ln(\frac{S_0}{K}) + (\mu - \frac{1}{2}\eta_i^2)T}{\eta_i \sqrt{T}} \right) \right]$$

where

$$\eta_i := \frac{V_i(T)}{\sqrt{T}} = \sqrt{\frac{\int_0^T \sigma_i^2(t)dt}{T}}$$

$\omega = 1$ for a call option and $\omega = -1$ for a put option. The option price formula expressed in proposition 3.1.2 gives rise to symmetric smiles in the implied volatility structure with a minimum at a strike price equal to the forward asset price $S_0 \exp(\mu T)$ [21]. The presence of a relative minimum is a major disadvantage when fitting asymmetric smiles (skews). To circumvent this problem, Brigo and Mercurio [21] derived alternative local volatility mixture

models. By adding a new parameter, α they were able to shift the original asset price dynamics to fit asymmetric skews. Another alternative was to allow the $p_t^i(\cdot)$ densities to remain lognormal but with different means. The mixture model with a shifted log normal distribution is now discussed.

3.1.1.2 Shifted mixture of lognormals

The original asset price, S_t, in this model is altered using an affine transformation i.e. the process described in equation 3.10 is shifted. The new asset price process becomes [18]:

$$A_t = A_0\alpha\exp(\mu t) + S_t \quad . \textbf{(3.11)}$$

where α is a real constant. This transformation is chosen in order to preserve the drift rate of the new asset price process, A_t, as μ [21]. Now, if the original asset price dynamics, under the forward measure, Q^T, were:

$$dS_t = \mu S_t dt + v(t, S_t)S_t dW_t . \quad \textbf{(3.12)}$$

where v is defined by the equation in proposition 3.1.1. Ito's formula can be applied to equation 3.12 to produce the dynamics of the new asset price process, also under the forward measure, Q^T [21].The calculations are as follows:

From (3.11):

$$S_t = A_t - A_0\alpha\exp(\mu t)$$

By Ito's formula:

$$dS_t = \mu S_t dt + v(S_t, t)S_t dW_t = dA_t - A_0\alpha\mu\exp(\mu t)$$

$$\therefore \mu\left(A_t - A_0\alpha\exp(\mu t)\right)dt + v(S_t, t)\left(A_t - A_0\alpha\exp(\mu t)\right)dW_t = dA_t - A_0\alpha\mu\exp(\mu t)$$

The equation above is simplified and solved for dA_t to produce the following stochastic differential equation:

$$dA_t = \mu A_t dt + v\left(A_t - A_0\alpha\exp(\mu t), t\right)(A_t - A_0\alpha\exp(\mu t))dW_t \quad \textbf{(3.13)}$$

where $v(y,t)$ is defined as before. The following proposition states the marginal density of the shifted mixture of lognormals.

Proposition 3.1.3 [18]:

If $v(y,t)$ is defined as in proposition 3.1.1, then the marginal density for the new shifted asset price process with dynamics given by equation 3.13, is given as:

$$p_t(y) = \sum_{i=1}^{N} \lambda_i \frac{1}{(y - A_0\alpha e^{\mu t})V_i(t)\sqrt{2\pi}} \exp\left\{-\frac{1}{2V_i^2(t)}\left[\ln\left(\frac{y - A_0\alpha e^{\mu t}}{A_0(1-\alpha)}\right) - \mu t + \frac{1}{2}V_i^2(t)\right]^2\right\}$$

With

$y - A_0\alpha e^{\mu t} > 0$

Some examples of the density function for A_T that are obtained for different values of α are shown in Figure 3.1.

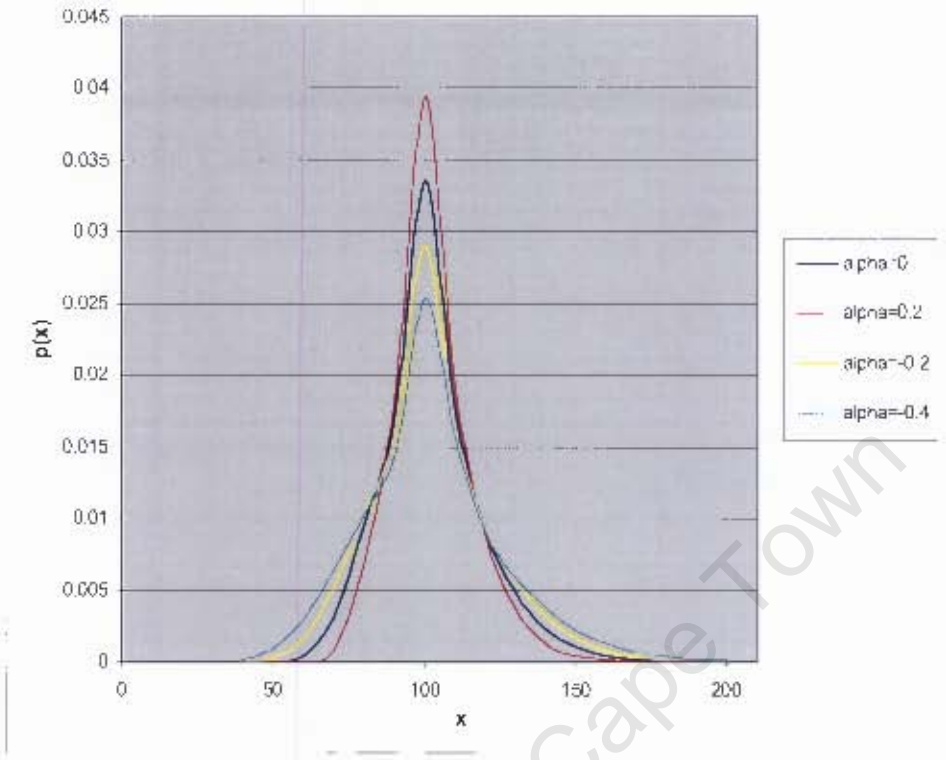


Figure 3.1: The density function of A_T for different values of α

Figure 3.1 illustrates how the shape of the density function is altered as different values for α are used. As α becomes more positive, the distribution becomes more peaked and its tails flatten. On the other hand, as α becomes more negative, the density function becomes less peaked while the tails become fatter.

The time zero price of a European call option, $EO(K,T,\omega=1)$ written on the asset with the dynamics in equation 3.12 can be calculated as:

$$P(0,T)\,E^T\left\{(A_T-K)^+\right\} \quad P(0,T)\,E^T\left\{(A_0\alpha\exp(\mu T)+S_T-K)^+\right\}$$

$$= P(0,T)E^T\left\{(S_T-[K-A_0\exp(\mu T)])^+\right\}$$

where E^T is the expectation under the Q^T forward measure. Proposition 3.1.4 gives the analytical formula of the European option price.

Proposition 3.1.4 [22]:

Consider a European option, $EO(K,T)$, written on the asset with the dynamics (3.13), with maturity T and strike K. Assuming that $K - A_0\alpha\exp(\mu T) > 0$, the value of the option at time zero is given as:

$$EO(K,T,\omega) =$$

$$\omega P(0,T)\sum_i^N \lambda_i \left[\Lambda_0 \exp(\mu T)\Phi\left(\omega \frac{\ln(\frac{\Lambda_0}{\kappa}) + (\mu + \frac{1}{2}\eta_i^2)T}{\eta_i\sqrt{T}} \right) - K\Phi\left(\omega \frac{\ln(\frac{\Lambda_0}{\kappa}) + (\mu - \frac{1}{2}\eta_i^2)T}{\eta_i\sqrt{T}} \right) \right]$$

Where:

$$\kappa = K - A_0\alpha\exp(\mu T)$$

$$\Lambda_0 = A_0(1-\alpha)$$

where η_i and ω are defined as before.

This option pricing formula gives rise to asymmetric volatility smiles provided that $\alpha \neq 0$. An alternative model that also leads to asymmetric smiles is based on a mixture of lognormal densities with different means. This model is now discussed.

3.1.1.3 Log Normal mixture model with different means

The diffusion process for each mixture component is now given as [19]:

$$dS_t^i = \mu_i(t)S_t^i dt + \sigma_i(t)S_t^i dW_t \quad . \qquad (3.14)$$

where σ_i's are defined as before and are μ_i are deterministic functions of time. The density of each process is still lognormal but it is now expressed as:

$$p_t^i(y) = \frac{1}{yV_i(t)\sqrt{2\pi}} \exp\left\{ -\frac{1}{2V_i^2(t)} \left[\ln(\frac{y}{S_0}) - M_i(t) + \frac{1}{2}V_i^2(t) \right]^2 \right\}$$

where $M_i(t) := \int_0^t \mu_i(u)du$

Furthermore, each μ_i must be chose such that [22]:

$$\sum_{i=1}^{N} \lambda_i e^{M_i(t)} = e^{\mu t} \quad . \qquad (3.15)$$

This is a no-arbitrage condition that will ensure that $p_t(y)$, is a proper density with mean equal to $S_0 e^{\mu t}$ [22]. Now, applying the Fokker-Planck equation once again to the S and S_i processes, the diffusion coefficient, $\Psi(y,t)$ can be extracted [22]. This coefficient ensures that the equation:

$$dS_t = \mu S_t dt + \Psi(S_t,t)S_t dW_t \,. \qquad (3.16)$$

has a solution and a density given by a mixture of lognormal densities.

The diffusion coefficient $\Psi(y,t)$ can be expressed as follows [19]:

$$\Psi(y,t) = \frac{\sum_{i=1}^{N} \lambda_i \sigma_i^2(t) p_t^i(y)}{\sum_{i=1}^{N} \lambda_i p_t^i(y)} + \frac{2S_0 \sum_{i=1}^{N} \lambda_i (\mu_i(t) - \mu) e^{M_i(t)} \Phi\left(\frac{\ln(\frac{y}{S_0}) - M_i(t) + \frac{1}{2} V_i^2(t)}{V_i(t)} \right)}{y^2 \sum_{i=1}^{N} \lambda_i p_t^i(y)}$$

The derivation can be found in [19]. The time zero price of a European option, $EO(K,T,\omega)$ written on the asset with the dynamics in equation 3.16 can also be written as a convex combination of Black-Scholes prices [19]. Its value at time zero is given as:

$$EO(K,T,\omega) =$$

$$\omega P(0,T) \sum_{i=1}^{N} \lambda_i \left[S_0 \exp(M_i(T)) \Phi\left(\omega \frac{\ln(\frac{S_0}{K}) + M_i(T) + \frac{1}{2}\eta_i^2 T}{\eta_i \sqrt{T}} \right) - K\Phi\left(\omega \frac{\ln(\frac{S_0}{K}) + M_i(T) - \frac{1}{2}\eta_i^2 T}{\eta_i \sqrt{T}} \right) \right]$$

where η_i and ω are defined as before.

3.1.2 Limitations of the models

Although local volatility mixture models are flexible enough to recover a large variety of implied volatility shapes, they do not model random fluctuations of asset prices such as crashes and rallies that take place over a period of time. Such fluctuations are modelled by jump diffusion models. These models are discussed in Section 3.2.

3.2 *Jump Diffusion models*

The Black-Scholes model fails to explain random fluctuations of asset prices such as crashes and rallies that take place over a period of time. In order to achieve better modeling of reality, jumps and crash components have been added to the Black-Scholes model to form jump diffusion models [43].

From a risk management perspective, jump diffusion models are important because they are able to explain discontinuous asset price jumps. The occurrences of events over a pre-defined period are most commonly modeled by Poisson processes [32]. It is this reason that makes them attractive for use in jump diffusion models. Different distributions were chosen by researchers to model the size of the jumps. Merton [17] used a log normal process to describe the distribution of the jump amplitudes. Kou [30], on the other hand, selected a log double exponential distribution for the jump amplitude process. In [31], a uniform distribution is used to describe the log of the jump amplitude process. These models:

- Lead to leptokurtosis in the distribution of underlying asset returns
- Give rise to a volatility skew

Only Merton and Kou's models, however, are able to provide analytical solutions for call and put options [30]. The log uniform model, on the other hand uses numerical methods such as Monte Carlo algorithms for option pricing. Its main advantage over the other two models is that it is able to produce fatter tails in its log return distribution. Its log return distribution, therefore, more closely resembles that of log returns of market data.

The double exponential jump diffusion model described in [30] is implemented in this thesis. It is now discussed.

3.2.1 The double exponential jump diffusion model

Unlike Merton's jump diffusion model which uses a lognormal process, this model uses a double exponential process for the log of the jump amplitudes. The jump times nevertheless, still correspond to the event times of Poisson processes. The model is able to produce:

- Higher peaks and fatter tails in the asset returns distributions than the normal distribution
- Analytical solutions for European vanilla option prices
- The volatility skew.

Two properties of the double exponential distribution that make this possible are [30]:

- The leptokurtic feature of the double exponential distribution
- Its memory less property which is inherited from the exponential distribution. This feature enables closed form solutions for various options to be derived.

Both properties are discussed in Sections 3.2.1.1 and 3.2.1.2.

The double exponential model essentially consists of two parts. The first part is a continuous process modeled by a Brownian motion. A Poisson process with the log of the jump amplitude modeled as a double exponential distribution constitutes the second part of the model [30]. Under the real world measure, P, the following SDE is used to model the dynamics of the asset price S_t , at time t:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW(t) + d(\sum_{i=1}^{N(t)} (V_i - 1)) . \quad \textbf{(3.17)}$$

$W(t)$ is a standard Wiener motion and $N(t)$ is a Poisson process with rate λ . In the model, the drift μ and volatility σ are assumed to remain constant. Additionally, $W(t)$, $N(t)$ and Υ are assumed to be independent. The set $\{V_i\}$

represents a sequence of independently and identically distributed random variables which are modeled such that $\Upsilon = \log(V)$ has an asymmetric double exponential distribution. The density of Υ can be expressed as follows [30]:

$$f_{\Upsilon}(y) = p\eta_1 e^{-\eta_1 y} I_{\{y \geq 0\}} + q\eta_2 e^{\eta_2 y} I_{\{y<0\}} \ . \qquad \textbf{(3.18)}$$

$$\eta_1 > 1 \ , \ \eta_2 > 0 \text{ and } p, q \geq 0, p+q = 1$$

p and q represent the probabilities of upward and downward jumps respectively. η_1 and η_2 on the other hand, represent parameters of the upward and downward jumps respectively. It must be noted that $Var(\Upsilon)$ gives a measure of the jump volatility. $E(\Upsilon)$ gives an indication of the jump size mean.

The relationship between Υ and V can also be expressed as follows:

$$\log(V) = \Upsilon \overset{d}{=} \begin{cases} \xi^+ & \text{with probability p} \\ -\xi^- & \text{with probability q} \end{cases}$$

where the symbol $\overset{d}{=}$ means equal in distribution [30]. ξ^+ and ξ^- represent exponential random variables. The means of the respective exponential random variables are described by the parameters $\frac{1}{\eta_1}$ and $\frac{1}{\eta_2}$. The means $E(\Upsilon)$, $E(V)$ and variance, $Var(\Upsilon)$ of the asymmetric double exponential distribution can therefore be expressed as:

$$E(\Upsilon) = \frac{p}{\eta_1} - \frac{q}{\eta_2}$$

$$Var(\Upsilon) = pq(\frac{1}{\eta_1} + \frac{1}{\eta_2})^2 + (\frac{p}{\eta_1^2} + \frac{q}{\eta_2^2})$$

and **(3.19)**

$$E(V) = E(e^{\Upsilon}) = q\frac{\eta_2}{\eta_2 + 1} + p\frac{\eta_1}{\eta_1 - 1}$$

$$\eta_1 > 1 \text{ , } \eta_2 > 0$$

Solving equation 3.14 gives rise to the following solution for the dynamics of the asset price [30]:

$$S_t = S_0 \exp\left\{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t)\right\} \prod_{i=1}^{N(t)} V_i$$

$$= S_0 \exp\left\{(\mu - \frac{1}{2}\sigma^2)t + \sigma W(t) + \sum_{i=1}^{N(t)} \Upsilon_i\right\}$$

The condition on η_1 expressed in equation 3.15 therefore ensures that both $E(V) < \infty$ and therefore $E(S_t) < \infty$. This condition prevents the upward movement of the stock rising to infinity. The leptokurtic feature of the distribution of the returns of the asset in this model is now discussed. Option pricing is discussed in Section 3.2.1.2.

3.2.1.1 Leptokurtic property

The leptokurtic feature of the distribution of returns of the asset of can be demonstrated as follows:

$$\frac{\Delta S_t}{S_t}$$

$$= \frac{S_{t+\Delta t} - S_t}{S_t}$$

$$= \frac{S_{t+\Delta t}}{S_t} - 1$$

$$= \exp\left\{(\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma(W(t+\Delta t) - W(t))\right\} \frac{\prod_{i=1}^{N(t+\Delta t)} \mathrm{V}_i}{\prod_{i=1}^{N(t)} \mathrm{V}_i} - 1$$

$$= \exp\left\{(\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma(W(t+\Delta t) - W(t)) + \sum_{i=1+N(t)}^{N(t+\Delta t)} \Upsilon_i\right\} - 1$$

Since $\Upsilon = \log(\mathrm{V})$ $\therefore$ $\mathrm{V} = \exp(\Upsilon)$

As Δt approaches zero, as in daily observations, the equation above can be approximated using the Maclaurin series expansion:

$$e^x \approx 1 + x + \frac{x^2}{2}$$

$$\therefore \frac{\Delta S_t}{S_t} \approx (\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma(W(t+\Delta t) - W(t)) + \sum_{i=N(t)+1}^{N(t+\Delta t)} \Upsilon_i + \frac{1}{2}\sigma^2(W(t+\Delta t) - W(t))^2$$

$$\approx \mu\Delta t + \sigma Z\sqrt{\Delta t} + \sum_{i=N(t)+1}^{N(t+\Delta t)} \Upsilon_i$$

Since $\frac{1}{2}\sigma^2(W(t+\Delta t) - W(t))^2 = 0$ [44]

and $W(t+\Delta t) - W(t) = Z\sqrt{\Delta t}$

where Z is a standard normal random variable

Now, the probability of a Poisson process having one jump in an interval Δt is $\lambda\Delta t$ and the probability of having more than one jump will involve computing higher order powers of $\lambda\Delta t$. If $\lambda\Delta t$ is assumed to be small, multiple jumps can be ignored and the following result is obtained [34]:

$$\sum_{i=1+N(t)}^{N(t+\Delta t)} \Upsilon_i = \begin{cases} \Upsilon_i \text{ with probability } \lambda\Delta t \\ 0 \text{ with probability } 1\text{-}\lambda\Delta t \end{cases}$$

Therefore:

$$\frac{\Delta S_t}{S_t} \approx \mu\Delta t + \sigma Z\sqrt{\Delta t} + B\Upsilon \qquad \textbf{(3.20)}$$

where B is a Bernoulli variable such that:

$$P(B=1) = \lambda\Delta t \text{ and } P(B=0) = 1 - \lambda\Delta t$$

An expression for the density function of the right hand side of equation 3.17 can be found in [30]. When the approximated density function is compared to a normal distribution with the same mean and variance, it is found to be more peaked around the mean and to have heavier tails [30].

3.2.1.2 Option pricing

Another reason for using the double exponential distribution in this jump diffusion model is because of the distribution's memory less property [30]. This property, which is inherited from the exponential distribution, gives rise to closed form solutions for option prices.

Definition 1 [29]:

"A variable, x, is said to be memory less with respect to t, if for s>0:

$$P(x > s+t \mid x > t) = P(x > s)$$

If $P(A \mid B) = \frac{P(A, B)}{P(B)}$ *then:*

$$P(x > s+t, x > t) = P(x > s)P(x > t)$$

Since $s > 0$ and $s+t > t$ *then:*

$$\begin{aligned} P(x > s+t, x > t) &= P(x > s+t) \\ &= P(x > s)P(x > t) \end{aligned}$$

The probability density function of the exponential distribution is defined as

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{for } x<0 \end{cases}$$

where λ is a parameter of the distribution.

From this, it can be verified that the exponential distribution has the memory less property as follows:

$$\mathrm{P}(\mathbf{x} > s+t) = e^{-\lambda(s+t)}$$

$$= e^{-\lambda s} e^{-\lambda t}$$

$$= \mathrm{P}(\mathbf{x} > s)\,\mathrm{P}(\mathbf{x} > t)$$

The memory less property is used in the jump diffusion model to decompose the sum of double exponential random variables.

As stated previously:

$$\Upsilon \overset{d}{=} \begin{cases} \xi^{+} & \text{with probability p} \\ -\xi^{-} & \text{with probability q} \end{cases}$$

where the symbol $\overset{d}{=}$ means equal in distribution [30]. ξ^{+} and ξ^{-} represent exponential random variables with rates η_1 and η_2. From the memory less property, the following can be derived:

$$\mathrm{P}\left(\xi^{+}-\xi^{-}<t \mid \xi^{+}>\xi^{-}\right)=1-\mathrm{P}\left(\xi^{+}-\xi^{-}>t \mid \xi^{+}>\xi^{-}\right)$$

$$=1-\mathrm{P}\left(\xi^{+}>t\right)$$

$$=\mathrm{P}\left(\xi^{+}<t\right)$$

And

$$\mathrm{P}\left(\xi^{+}-\xi^{-}<t \mid \xi^{+}<\xi^{-}\right)=\mathrm{P}\left(\xi^{-}>\xi^{+}-t \mid \xi^{-}>\xi^{+}\right)$$

$$=\mathrm{P}\left(\xi^{-}>-t\right)$$

$$=\mathrm{P}\left(-\xi^{-}<t\right)$$

These results yield $\left(\xi^{+}-\xi^{-} \mid \xi^{+}>\xi^{-}\right) \overset{d}{=} \xi^{+}$ and $\left(\xi^{+}-\xi^{-} \mid \xi^{+}<\xi^{-}\right) \overset{d}{=} -\xi^{-}$ and therefore the following [30]:

$$\xi^{+}-\xi^{-} \overset{d}{=} \begin{cases} \xi^{+} & \text{with probability } \dfrac{\eta_2}{\eta_1+\eta_2} \\ -\xi^{-} & \text{with probability } \dfrac{\eta_1}{\eta_1+\eta_2} \end{cases}$$

Since $\mathrm{P}\left(\xi^{+}>\xi^{-}\right)=\dfrac{\eta_2}{\eta_1+\eta_2}$ and $\mathrm{P}\left(\xi^{+}<\xi^{-}\right)=\dfrac{\eta_1}{\eta_1+\eta_2}$.

The results thus far can be extended to produce the following [30]:

$$\sum_{i=1}^{n} \Upsilon_i \overset{d}{=} \begin{cases} \sum_{i=1}^{k} \xi_i^{+} & \text{with probability } P_{n,k},\ \text{k=1,2,....,n} \\ -\sum_{i=1}^{k} \xi_i^{-} & \text{with probability } Q_{n,k},\ \text{k=1,2,....,n} \end{cases}$$

Where $P_{n,k}$ and $Q_{n,k}$ are given by

$$P_{n,k} = \sum_{i=k}^{n-1} \binom{n-k-1}{i-k} \binom{n}{i} \left(\frac{\eta_1}{\eta_1+\eta_2}\right)^{i-k} \left(\frac{\eta_2}{\eta_1+\eta_2}\right)^{n-i} p^i q^{n-i}$$

$$Q_{n,k} = \sum_{i=k}^{n-1} \binom{n-k-1}{i-k} \binom{n}{i} \left(\frac{\eta_1}{\eta_1+\eta_2}\right)^{n-i} \left(\frac{\eta_2}{\eta_1+\eta_2}\right)^{i-k} p^{n-i} q^i$$

Where $1 \le k \le n-1$, $P_{n,n} = p^n$, $Q_{n,n} = q^n$

The sum of exponential random variables is equal to a gamma random variable [30]. The result above implies that the sum of i.i.d. (independently and identically distributed) double exponential random variables is equal in distribution to a single mixed gamma random variable [34]. This result simplifies the derivation of a closed form solution options because there is now only one random variable to work with rather than a sum of variables.

The price of a vanilla call option is derived as follows:

$$\begin{aligned} \Psi_c(t) &= \mathrm{E}\left(e^{-rt} * payoff\right) \\ &= \mathrm{E}\left(e^{-rt}[S_T - K]^+\right) \end{aligned}$$

In order to compute the price of the option, it is necessary to derive the cumulative density function of the asset price process in equation 3.16. To do this, the distribution of the sum of normal and double exponential random variables must be investigated. The distribution can be expressed in closed form as an Hh function [32]. This function is defined as follows:

$$Hh_n(x) = \int_x^{\infty} Hh_{n-1}(y)dy = \frac{1}{n!}\int_x^{\infty}(t-x)^n e^{-\frac{t^2}{2}}dt \geq 0$$

for n = 0, 1 , 2

where:

$$Hh_{-1}(x) = e^{-\frac{x^2}{2}}$$

$$Hh_0(x) = \sqrt{2\pi}\Phi(-x)$$

$\Phi(x)$ is a cumulative normal distribution

The Hh functions can also be expressed as a three-term recursion [32]:

$$nHh_n(x) = Hh_{n-2}(x) - xHh_{n-1}(x), \qquad n \geq 1$$

Figure 3.2 illustrates the Hh function for n =1, 3 and 5

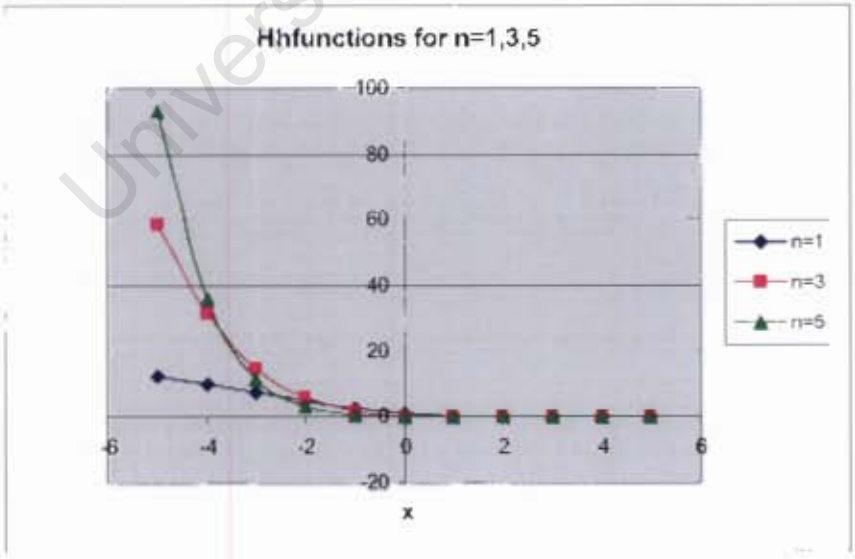


Figure 3.2: The Hh function for n=1, 3 and 5

Hh functions for three different values for n are plotted in Figure 3.2. For x < 0, as the value for n is increased, the gradient of the graph becomes steeper. The Hh function is used to compute the I_n integral. This integral is used in the option pricing formula. I_n is computed as follows[30]:

$$I_n(c;\alpha,\beta,\delta) := \int_c^{\infty} e^{\alpha x} Hh_n(\beta x - \delta)dx \ , \ n \geq 0$$

where α, β and c represent arbitrary constants. This integral can be further simplified to equations B7 and B8 in [30]. In order to give closed forms for European call and put options, for probability measure P , define:

$$\Gamma(\mu,\sigma,\lambda,p,\eta_1,\eta_2;a,T) = \mathrm{P}\{(Z(T) \geq a\}$$

Where:

$$Z(t) = \mu t + \sigma W(t) + \sum_{i=1}^{N(t)} \Upsilon_i$$

The pricing formula for a European call option Ψ_c can be expressed in terms of the Γ function as follows:

$$\Psi_c(0) = S_0\ \Gamma(r + \frac{1}{2}\sigma^2 - \lambda\varsigma\ ,\ \sigma\ ,\ \tilde{\lambda}\ ,\ \tilde{p}\ ,\ \tilde{\eta}_1\ ,\ \tilde{\eta}_2\ ;\ \log(\frac{K}{S_0})\ ,\ T)$$

$$- Ke^{-rT}\ \Gamma(r - \frac{1}{2}\sigma^2 - \lambda\varsigma\ ,\ \sigma\ ,\ \lambda\ ,\ p\ ,\ \eta_1,\ \eta_2\ ;\ \log(\frac{K}{S_0})\ ,\ T)$$

where

$$\tilde{p} = \frac{p}{1+\varsigma}\frac{\eta_1}{\eta - 1}$$

$$\tilde{\eta}_1 = \eta_1 - 1$$

$$\tilde{\eta}_2 = \eta_2 + 1$$

$$\tilde{\lambda} = \lambda(\varsigma + 1)$$

$$\varsigma = \frac{p\eta_1}{\eta_1 - 1} + \frac{q\eta_2}{\eta_2 - 1} - 1$$

The Γ function is derived from the sum of Hh functions and the I_n integral. An explicit formula for Γ can be found in [30]. In the limiting case, i.e. as $\lambda \to 0$, the jump diffusion model reverts back to the Black-Scholes model and call option prices convergence to Black-Scholes prices. For a proof see [30].

Using put-call parity, the price of a put can be derived as follows:

$$\Psi_P(0) - \Psi_C(0) = Ke^{-rT} - S_0$$

3.2.2 Shortcomings of the model

Despite the attractive features presented by the double exponential jump diffusion model, some of its shortcomings have been documented in [30], [34], and [35]. They include:

- Compared to other jump diffusion models, this model requires the estimation of more parameters when calibrating. As a result more data is needed for fitting.
- Model calibration is more accurate when more liquid options are used.
- Tails of log double exponential jump decay exponentially whereas financial market log return data displays thicker tails
- Infinite jump domain- implies that jumps are not bounded. This is unrealistic because jumps are bounded.

Each model described in Chapters 2 and 3 was implemented. The data and methodology used in this thesis are now discussed.

4 Methodology

The procedures and algorithms that were used to implement the four models that were discussed in the previous chapters are now outlined.

Only European vanilla option prices were computed in this thesis. In order to produce volatility skews from these prices, the following assumption was made:

- Interest rates for a single maturity are constant and are equal to the risk-free rate, i.e. $r > 0$.

As previously mentioned, parameters of each model were determined in one of two ways:

1. Estimation from the historical prices of the underlying stock

   Parameters in the Edgeworth and GARCH models were estimated using historical prices.

2. Calibration to the current prices of liquid options

   The local volatility mixture and jump diffusion models were calibrated to market implied volatility data.

4.1 The Edgeworth model

As highlighted in Chapter 2, Rubinstein's option valuation method consists of two steps. These are:

- Generating an Edgeworth distribution
- Option pricing using Edgeworth probabilities

The data used to price options via an Edgeworth expansion consisted of daily closing prices of the Johannesburg Stock Exchange (JSE) TOPI40 index over a five year period ranging from 17/01/2000 -16/09/2005. This index is also known as the JSE All Share Index (ALSI) 40.

The JSE TOPI40 Index is a market value weighted index that is based on a portfolio of the top 40 shares (by market capitalisation) traded on the Johannesburg Stock Exchange. The weight of each stock in the index is therefore proportional to its market capitalisation.

Details of the modelling procedure are now discussed.

4.1.1 Generating an Edgeworth distribution

The spreadsheet layout used to generate an Edgeworth distribution with known skewness and kurtosis values is illustrated in Figure 4.1.

Generating an Edgeworth distribution

Mean	0.0081837
Variance	0.8755851
Skew	-0.3202023
Kurtosis	4.1478218

Steps	16
Skew- e	-0.386731
Kurt- k	6

Σf(xj)	Σf'(xj)
0.9845048	1

J	X(j)	b(xj)	C-Edge factors	f(xj)	f'(xj)	x'(j)
0	-4	1.52588E-05	26.72	0.0004078	0.0004142	-4.2835
1	-3.5	0.000244141	13.29	0.0032443	0.0032954	-3.74915
2	-3	0.001831055	5.71	0.0104567	0.0105213	-3.21481
3	-2.5	0.008544922	1.94	0.0165522	0.0168127	-2.68047
4	-2	0.027770996	0.48	0.0133595	0.0135698	-2.14612
5	-1.5	0.066650391	0.29	0.0195208	0.019828	-1.61178
6	-1	0.122192383	0.65	0.0799536	0.081212	-1.07743
7	-0.5	0.174560547	1.10	0.1914898	0.1945037	-0.54309
8	0	0.196380615	1.34	0.2633044	0.268058	-0.00875
9	0.5	0.174560547	1.27	0.2224309	0.2259318	0.525598
10	1	0.122192383	0.91	0.1114573	0.1132115	1.059942
11	1.5	0.066650391	0.44	0.0291867	0.029646	1.594286
12	2	0.027770996	0.22	0.0061996	0.0062972	2.12863
13	2.5	0.008544922	0.89	0.0075023	0.0077219	2.662974
14	3	0.001831055	3.39	0.006208	0.0063057	3.197317
15	3.5	0.000244141	9.12	0.0022254	0.0022604	3.731661
16	4	1.52588E-05	20.02	0.0003055	0.0003103	4.266005

Figure 4.1: Generating an Edgeworth distribution with known skewness and kurtosis

Figure 4.1 illustrates the generation of an Edgeworth distribution with a kurtosis of 6 and a skewness of -0.386731 using a 16 step binomial process. Skewness and kurtosis values were obtained from the JSE TOPI40 index returns data. This data was computed from the daily closing prices of the JSE TOPI40 index.

Binomial variables and their probabilities are represented by $x(j)$ and $b(xj)$ respectively. C-Edge factors are the Edgeworth expansion terms. The standardised probability distribution and its adjustment are displayed in columns $f(xj)$ and $f'(xj)$ respectively. The mean and variance of the new probability distribution $f'(xj)$ were also calculated and displayed in the cells labelled mean and variance. They were then used to transform the random variables x (j) into Edgeworth variables, $x'(j)$.

Once Edgeworth variables were computed, they were used for pricing options on the index. Option pricing procedures are discussed in the following section.

4.1.2 Option valuation

The second step in Rubinstein's valuation method involved generating a distribution of asset prices from the adapted binomial distribution and calculating option prices via the risk- neutral probabilities. This process was automated via the use of a coded subroutine. The inputs required to compute option values are illustrated in Table 4.1.

Table 4.1: Table of inputs used to produce option prices using the Edgeworth expansion

Today ' s index level S	13000
Risk-free rate r (%)	7.6
Dividend yield q (%)	0
ATM Volatility σ (%)	12
Strike Price K	9000
Time to maturity (T, years)	0.1890
ξ (skewness)	0.00
κ (kurtosis)	3.00
Nsteps	100
Call/Put (1/-1)	1

Table 4.1 displays the inputs that a user must specify in order to calculate the price of a call option. The current index level is 13000. The entry Nsteps was used to quantify the number of steps for the binomial process. Skewness and kurtosis values were adjusted so that the resulting implied volatility skew closely matched that observed in the market. When choosing skewness and kurtosis values, the table of permissible (ξ,κ) pairs found in Appendix A was consulted to avoid negative non- unimodal distributions.

4.1.3 Description of algorithms and user defined functions

The *EDVAL* subroutine was used to generate an Edgeworth distribution. It was also used to produce a put or call option price as an output. The inputs that were required to price an option using the Edgeworth model are illustrated in Table 4.1.

Additional functions were also written to test the resulting Edgeworth probability distribution. These functions, which are called within the EDVAL subroutine, were used to determine whether the Edgeworth probability density function had negative entries and more than one mode. In the event that this was the case, the Edgeworth expansion was replaced by the Gram-Charlier expansion. If the Gram-Charlier expansion method still gave rise to a distribution that was not unimodal and had negative entries, the *EDVAL* subroutine would produce as an output a value of -1 .This result would indicate that different skewness and kurtosis inputs values are required. The user must therefore specify different input values that are close to the permissible (ξ, κ) pairs found in Table A1 in Appendix A.

The algorithm used in the modelling process is outlined in Figure 4.2:

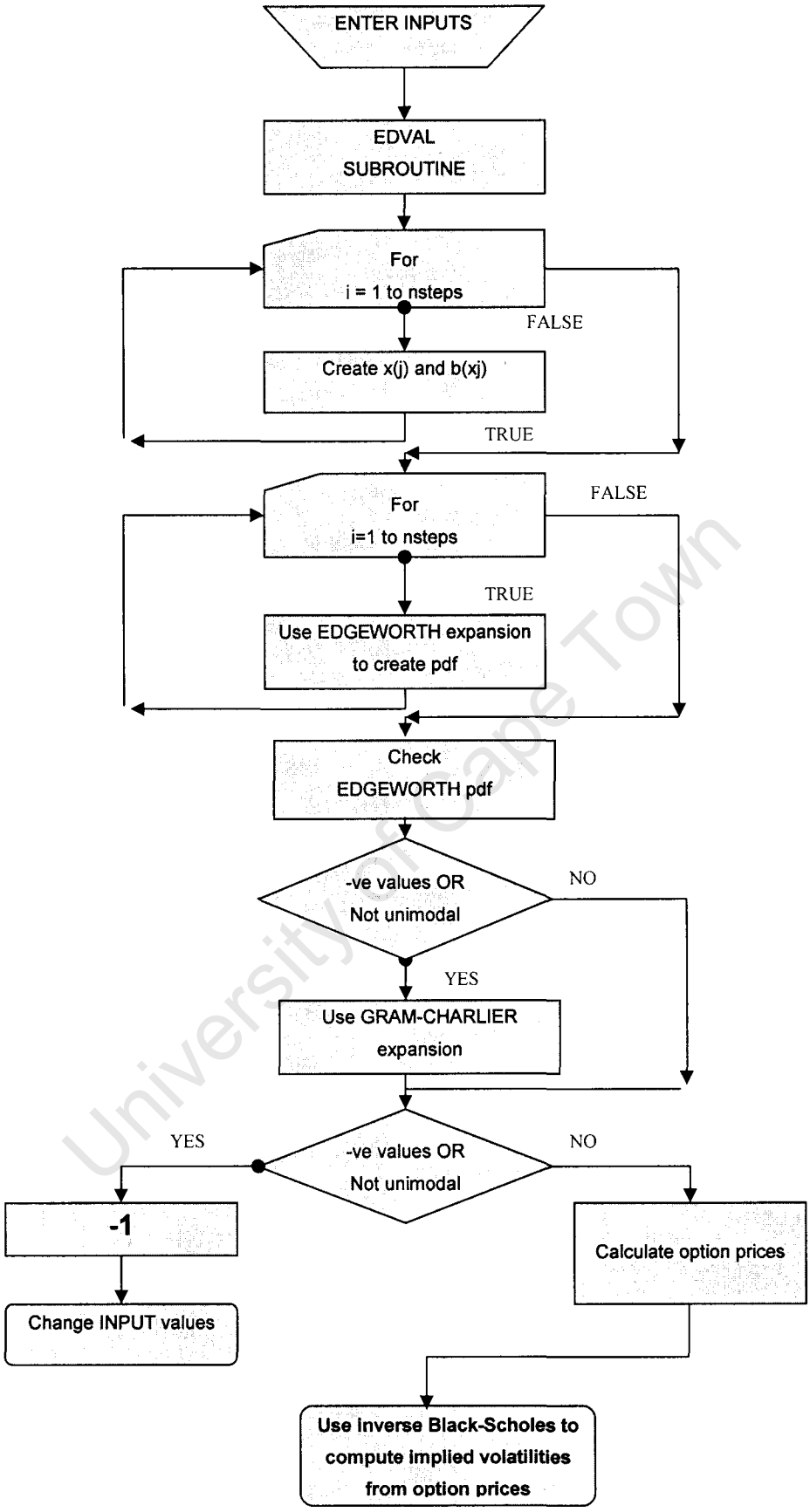


Figure 4.2: Flow chart of algorithm to price options using the Edgeworth Expansion

4.2 The asymmetric normal mixture GARCH (1, 1) model

In research such as [41], it has been concluded that normal mixture models with two variance components are sufficient to avoid estimation problems. As a result, a normal mixture GARCH (1, 1) model was fitted with only two variance components in this thesis. Furthermore, in order to give rise to leptokurtic distributions, a normal mixture AGARCH (1, 1) model with zero means in the mixture was implemented. This model shall now be referred to as the NM(2) AGARCH model.

As with the Edgeworth model, the data used was based on daily closing prices of the JSE TOPI40 over the period 17/01/2000 -16/09/2005. This data was used for parameter estimation and option pricing.

The procedures used to estimate the parameters of an NM (2) AGARCH model with zero means in the mixture are now discussed. Option pricing procedures are also described in Section 4.2.2.

4.2.1 Model parameter estimation

In order to give rise to skewed leptokurtic distributions ,The NM(2) AGARCH model ,with zero mean, was fitted i.e. the error term in this model was assumed to be a probability weighted average of two normal density functions. Each density function therefore had a mean of zero and variance modelled by the following AGARCH process:

$$\sigma_{it}^2 = \omega_i + \alpha_i(\varepsilon_{t-1} - \lambda_i)^2 + \beta_i\sigma_{it-1}^2 \quad \text{for } i = 1 \text{ and } 2$$

The overall conditional variance in the NM (2) AGARCH model was therefore:

$$\sigma_t^2 = p\sigma_{1t}^2 + (1-p)\sigma_{2t}^2$$

Nine parameters had to be estimated in this model. These parameters were estimated via maximum likelihood methods. The overall likelihood that was maximised in order to obtain optimal parameter values was:

$$L(\theta \mid \varepsilon) = \sum_{t=1}^{T} \ln\left[p\varphi_1(\varepsilon_t) + (1-p)\varphi_2(\varepsilon_t)\right]$$

Solver in excel, was used to maximise this expression. The function was used together with the following constraints:

1. $\omega_1, \omega_2 > 0$
2. $\alpha_1, \alpha_2 > 0$
3. $m = \dfrac{p(\omega_1 + \alpha_1 \lambda_1^{\,2})}{1-\beta_1} + \dfrac{(1-p)(\omega_2 + \alpha_2 \lambda_2^{\,2})}{1-\beta_2} > 0$
4. $n = \dfrac{p(1-\alpha_1-\beta_1)}{1-\beta_1} + \dfrac{(1-p)(1-\alpha_2-\beta_2)}{1-\beta_2} > 0$
5. $\omega_1 + \alpha_1\left(\dfrac{m}{n} + \lambda_1^2\right) > 0$
6. $\omega_2 + \alpha_2\left(\dfrac{m}{n} + \lambda_2^2\right) > 0$
7. $0 < p < 1$

These constraints are explained in Section 2.2. The parameters estimated were then used to calculate option prices and implied volatilities. For simplicity, these parameters were assumed to remain constant over an option's lifetime. The option pricing process is now discussed.

4.2.2 Option pricing

Once the NM (2) AGARCH model parameters were estimated, the Black-Scholes formula, with volatility input given as the square root of the expected annualised variance rates over the life of the options was used to calculate two individual option prices. The NM (2) AGARCH option price was then calculated as a weighted sum of the individual option prices. i.e.:

$$NM\_Option\_Price = p * Option\_Price\_1 + (1 - p) * Option\_Price\_2$$

The volatility skew was obtained by using the inverse Black-Scholes formula. The inputs required to compute option values are illustrated in Table 4.2.

Table 4.2: Table of inputs used to produce option prices using the NM (2) AGARCH model

Today's date	2005/01/17
Option Expiry Date	2005/09/16
Today's Index level	12000
Expected annualised volatility 1 (%)	10.8
Expected annualised volatility 2 (%)	9.2
Dividend yield (%)	0
Risk Free Rate (%)	7.6
Call/ Put (1/-1)	1

Table 4.2 shows the inputs required to calculate the price of a call option on the JSE TOPI40 index. The expected annualised volatilities in the table were calculated from the expected variance rate of the option over its lifetime.

4.2.3 Description of algorithms and user defined functions

User defined functions and subroutines were written for both parameter estimation and option pricing. The NMGARCH subroutine was used to estimate the model parameters from the historical closing prices of the JSE TOPI40. The NM_CALL_PRICE subroutine was used to produce put or call option prices.

4.3 *Local volatility mixture models*

The three local volatility models that were investigated are:

- The lognormal mixture model

  This model consists of a mixture of lognormal distributions. Each mixture is assumed to have the same constant mean denoted by μ.

- The shifted mixture model of lognormal distributions

  Once again each mixture in this model is assumed to have a lognormal distribution with the same constant mean μ. The model however is obtained by shifting the previous asset price dynamics.

- The lognormal mixture model with different means

  Each mixture in this model is assumed to have a lognormal distribution with a different mean, μ_i.

Data containing call option prices, in the form of implied volatilities, on the JSE TOPI40 Index were used to calibrate these models. The JSE TOPI40 index option was chosen for modelling because its underlying index reflects the performance of the South African ordinary stock exchange and is thus representative of the entire stock market. Also, the JSE TOPI40 index option is one of the most actively traded options on the JSE and is therefore sufficiently liquid to have a skew built for it.

The procedures used to calibrate these models to market data are now discussed.

4.3.1 Calibration to market data

Calibration of the local volatility mixture models was based on obtaining model parameters that minimised the sum of the squared errors between the market and model option prices. The model parameters that were estimated were λ_i and η_i , for i= 1,...N, where N represents the number of mixtures in the model. For the shifted lognormal mixture model, an additional parameter i.e. α had to be estimated. $M_i(T)$ parameters for the lognormal mixture model with different means also had to be estimated.

The excel / vba *solver* function was used to obtain parameter estimates. *Solver* uses the Generalized Reduced Gradient (GRG2) nonlinear optimisation code. Although this function was easy to use, convergence to the optimum solution was very slow and sensitive to starting values. Different answers were obtained depending on the initial guess on the parameter values. As a result, some starting values, particularly the alpha parameter in the shifted distribution, had to be adjusted continuously to ensure that the resulting skew produced by the model, closely resembled the market skew.

It is suggested however, that future work on this topic be extended to develop or implement an optimisation algorithm that is fast and reliable. In particular, it is recommended in [21] that a global search algorithm should be used with a few parameters and the search must be refined using a local algorithm for the last solution. Although the global search is accurate, it is very slow, particularly if the number of mixtures (and therefore parameters) is large .The local algorithm, on the other hand, will speed up calculations but it is extremely sensitive to starting values. A combination of the two algorithms, as concluded in [21], should produce optimum results.

The least squares fitting method used to obtain parameters for the local volatility models was subject to the following constraints:

1. The mixing weights had to be positive and their sum equal to one. i.e.: $\lambda_i > 0$ and $\sum_{i=1}^{N} \lambda_i = 1$.
2. Model calibration also depended upon ensuring that each $\eta_i > 0$.
3. An additional constraint was imposed on the α parameter. For each traded strike, K, it was essential to ensure that $K > A_0 \alpha \exp(\mu T)$. This condition resulted in α being restricted to the following values: $\alpha < \frac{K}{A_0 \exp(\mu T)}$.
4. $M_i(T)$ parameters also had to be chosen such that the following no-arbitrage condition was satisfied: $\sum_{i=1}^{N} \lambda_i e^{M_i(T)} = e^{\mu T}$

In this thesis the only value for N investigated was 3 i.e. the index density had a mixture of 3 lognormal densities. Mixture models where N=2 and N=3 have been used in empirical work in [21] and [31]. In these studies, mixtures of 2 or 3 densities have been sufficient to produce accurate calibration results.

4.3.2 Description of algorithms and user defined functions

User defined functions and subroutines used to calibrate the models were written. The algorithm based on these functions and subroutines is presented in the Figure 4.3.

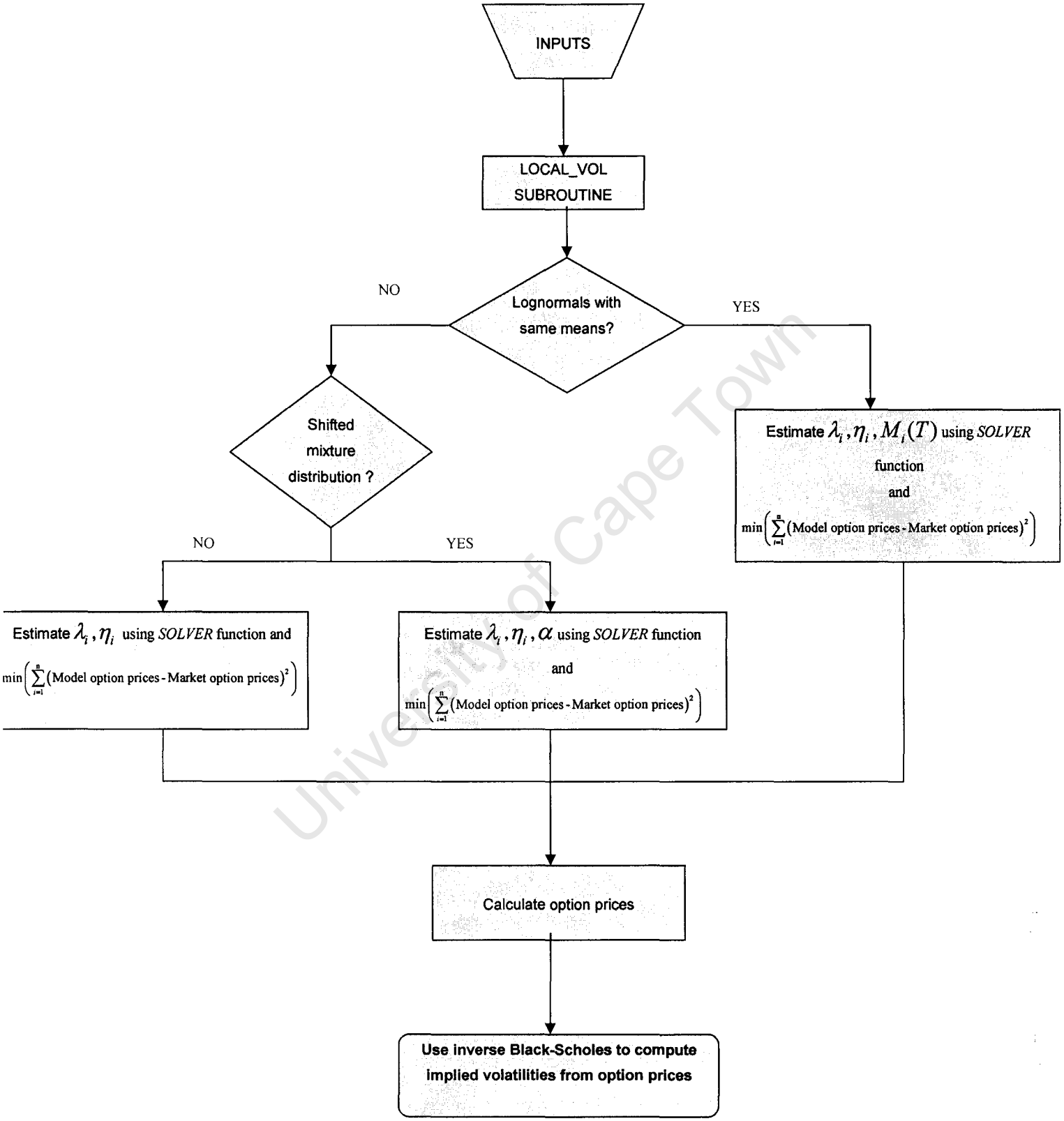


Figure 4.3: Flow chart illustrating local volatility model calibration

4.4 The double exponential jump diffusion model

The data described in Section 4.3 was used to calibrate the jump diffusion model. Procedures used to calibrate the jump diffusion model are now discussed.

4.4.1 Calibration to market data

Once again, model calibration was based on finding jump diffusion parameters that minimised the sum of the squared errors between the market and the model option prices. The solver function in excel was used to achieve this minimisation. Constraints that were applied to the parameters in this optimisation problem were:

1. The jump rate, $\lambda > 0$
2. $\eta_1 > 1$ and $\eta_2 > 0$.Reasons for these particular constraints are explained in Section 3.2.1.
3. The probability of an upward jump , $0 < p < 1$
4. The volatility of the asset price, $\sigma > 0$

Due to limitations imposed by Microsoft Excel, it was found that the maximum permissible value for η_1 was 300. The value for η_2, on the other hand, could not exceed 90.

4.4.2 Description of algorithms and user defined functions

The algorithm used to calibrate the model included:

- Creating Hh functions
- Computing I_n integrals
- Calculating option prices via the Γ function

User defined functions and subroutines to implement the above algorithm were written.

4.5 *Measuring model performances*

The aim of parameter estimation via historical data was to observe the implied volatility skew produced for the estimated parameters. A sensitivity analysis was used to determine the sensitivity of the model to variations of its parameters, in the vicinity of their best estimates.

In the case of calibration to market implied volatility data, both subjective and objective measures were used to evaluate the models. Subjective measures involved visually comparing the market and model skew. It was, however, difficult to determine how efficient a model is in reproducing the market skew using this method. Additionally assigning uncertainty or levels of reliability to parameter values was almost impossible. Objective measures on the other hand provided a more reliable means of comparing the fitting quality of the different models. The objective measures that were calculated are presented in Table 4.3.

Table 4.3 : Objective measures used to compare calibration results

Name	Description	Formula
ABSERR	Max Absolute Error	$Max\left\|g_{market} - g(\theta)\right\|$
ABSMEAN	Mean Absolute Error	$\frac{1}{n}\sum_{i=1}^{n}\left\|g_{market} - g(\theta)\right\|$
SSE	Sum of Errors Squared	$\sum_{i=1}^{n}\left(g_{market} - g(\theta)\right)^2$
RMSE	Root Mean Square Errors	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(g_{market} - g(\theta)\right)^2}$

In Table 4.3, market and model implied volatilities for n data points are represented by g_{market} and $g(\theta)$ respectively. θ is the vector of the model parameters.

5 Results and analysis

As discussed in Chapter 5, parameters of each of the four models investigated were determined in one of two ways i.e.:

- Estimation from the historical prices of the underlying index (Edgeworth and GARCH models).
- Calibration to the current prices of liquid options (Local and Jump diffusion models).

The results of these procedures are now presented in this section.

5.1 Results of estimation from historical prices

Sections 4.1 and 4.2 describe the methodology used to implement the Edgeworth and AGARCH models. The results from these two models are presented below.

5.1.1 The Edgeworth model results

Skewness and kurtosis values were obtained from the JSE TOPI40 index returns data. They were calculated to be as follows:

ξ (Skewness)	-0.13813
κ (Kurtosis)	3.4722

These values together with those displayed in Table 5.1 were used to calculate option prices.

Table 5.1: Inputs used to produce option prices using the Edgeworth expansion

Today's index level	12000
Risk-free rate (%)	7.5
Dividend yield (%)	0
Time to maturity (years)	0.1890
ATM volatility (%)	12
ξ (Skewness)	-0.13813
κ (Kurtosis)	3.4722
Nsteps	100
Call/Put	1

The resulting Edgeworth distribution using a 16 step process is displayed in Figure 5.1. It is compared to a standard normal distribution with kurtosis 3 and skewness of 0.

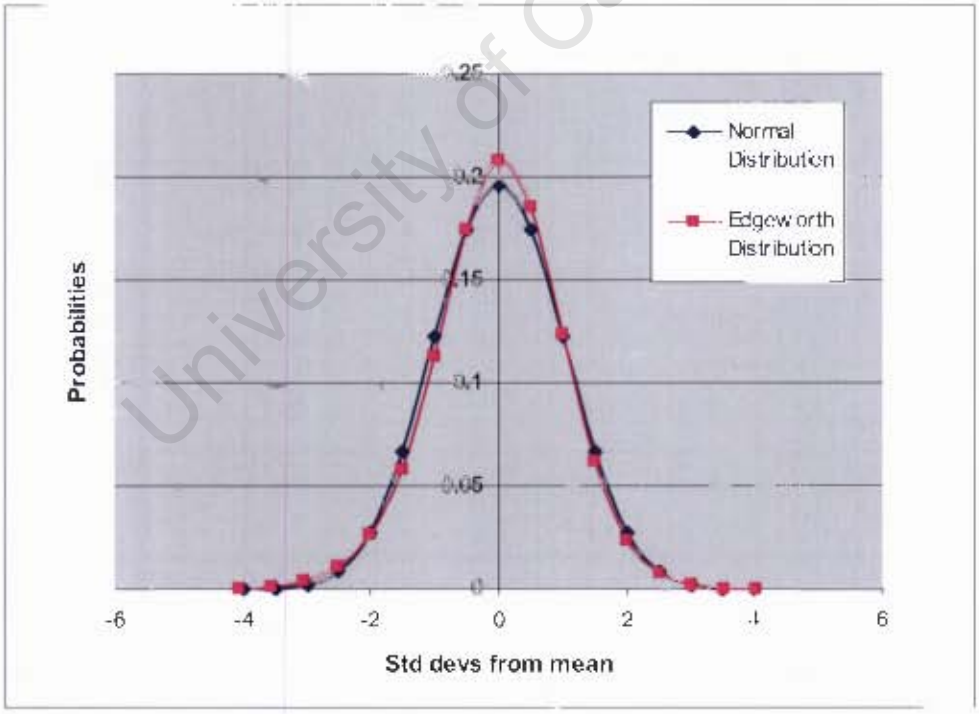


Figure 5.1: An Edgeworth distribution with kurtosis 3.4722 is compared to a standard normal distribution.

Figure 5.1 shows how an Edgeworth distribution increases the probabilities of extreme events (i.e. more than 2 standard deviations away from the mean) relative to a standard normal distribution. The Edgeworth distribution also has a slightly heavier left tail than the normal distribution. This implies that higher option prices and consequently higher implied volatilities are expected for low strike options. The resulting Edgeworth and Black-Scholes model implied volatilities are displayed in Figure 5.2.

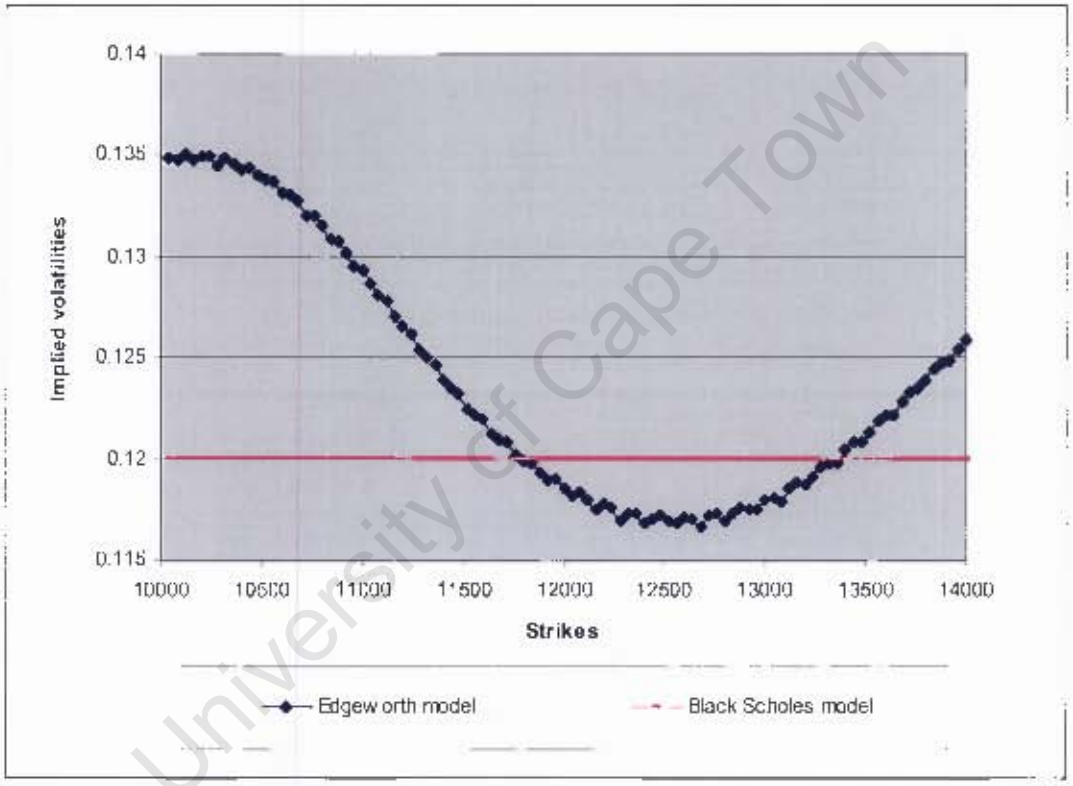
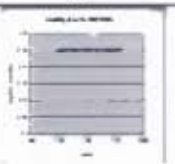
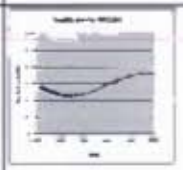
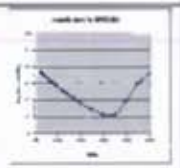
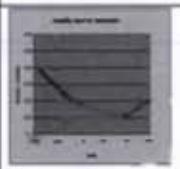
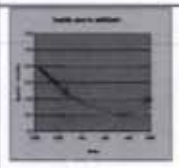
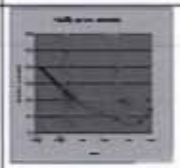
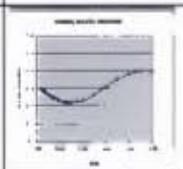
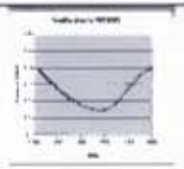
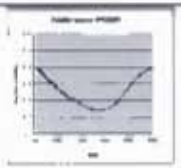
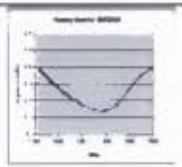
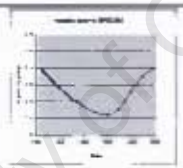
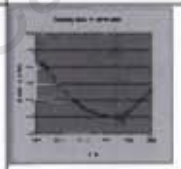
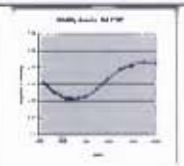
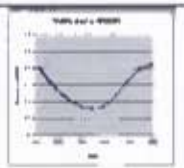
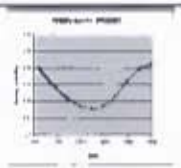
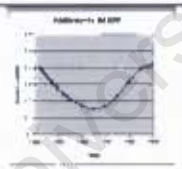
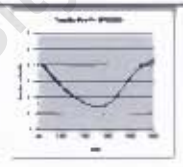
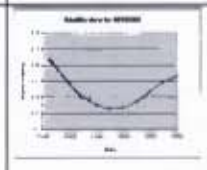
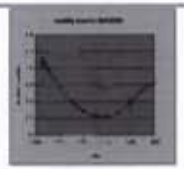
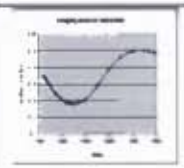
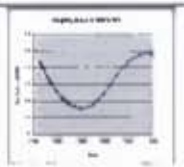
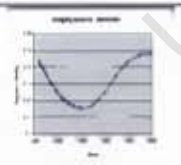
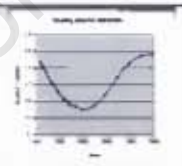
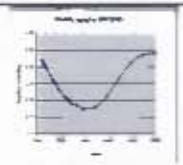
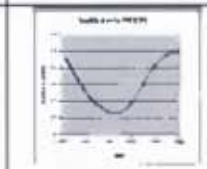
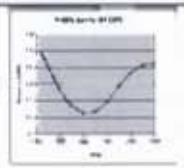
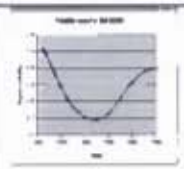
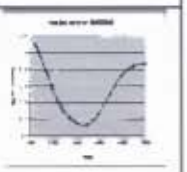


Figure 5.2: Implied volatilities produced from the inputs displayed in Table 5.1

The Black-Scholes model produces a constant implied volatility for the strike prices shown in Figure 5.2. On the other hand, a skew shape is produced by the Edgeworth model. Options with low strike prices (i.e. strikes less than the spot price of the underlying) are observed to have higher implied volatilities than options with higher strikes.

Table 5.2 Implied volatility shapes for different skewness (horizontal scale) and kurtosis (vertical scale) values

Skewness

	0	-0.38	-0.385	-0.39	-0.42	-0.5	-0.55	-0.7	-0.8
3		N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3.4						N/A	N/A	N/A	N/A
3.6							N/A	N/A	N/A
3.8								N/A	N/A
4.4									

The highlighted charts in Table 5.2 represent implied volatility shapes that closely resemble volatility skews observed in equity markets. The cells marked N/A represent skewness and kurtosis values that were not permissible. Table 5.2 illustrates that the volatility skew shape is sensitive to changes in skewness and kurtosis. For example, when the kurtosis is 3.4, a 1.3 % change in the skewness (from a value of -0.38 to -0.385) causes a significant change in the volatility skew shape.

5.1.2 The asymmetric normal mixture GARCH (1, 1) model results

The estimated model parameters are presented in Table 5.3.

Table 5.3 : NM (2) AGARCH parameters estimated for period 17/01/2000 -16/09/2005

ω_1	4.56E-06
α_1	0.098251839
β_1	0.756848786
λ_1	1.199429706
p	0.120001974
ω_2	3.94594E-06
α_2	0.095454132
β_2	0.845062272
λ_2	0.629839191

Substituting the above parameters values suggests that the time series of the variance process of each normal distribution is:

$$\sigma_{1t}^2 = 0.00000456 + 0.09825(\varepsilon_{t-1} - 1.1994)^2 + 0.7568\sigma_{1t-1}^2$$

$$\sigma_{2t}^2 = 0.000003945 + 0.09545(\varepsilon_{t-1} - 0.6298)^2 + 0.8451\sigma_{2t-1}^2$$

The inputs shown in Table 5.4 below were then used to compute option prices for various strike prices.

Table 5.4: Inputs used to calculate option prices using the NM (2) AGARCH model

Today's date	2005/01/17
Option Expiry Date	2005/09/16
Today's index level	11439
Expected annualised volatility 1 (%)	10.8
Expected annualised volatility 2 (%)	9.8
Dividend yield (%)	0
Risk Free Rate (%)	7.5
Call / Put (1/-1)	1

Table 5.4 above shows the inputs required to price a call option with a maturity of 0.66 years. The expected annualised volatilities in the table were calculated from the expected variance rate of the option over its lifetime. The call option prices that were computed from the inputs above were then used to calculate implied volatilities. Figure 5.3 illustrates the volatility skew that is produced.

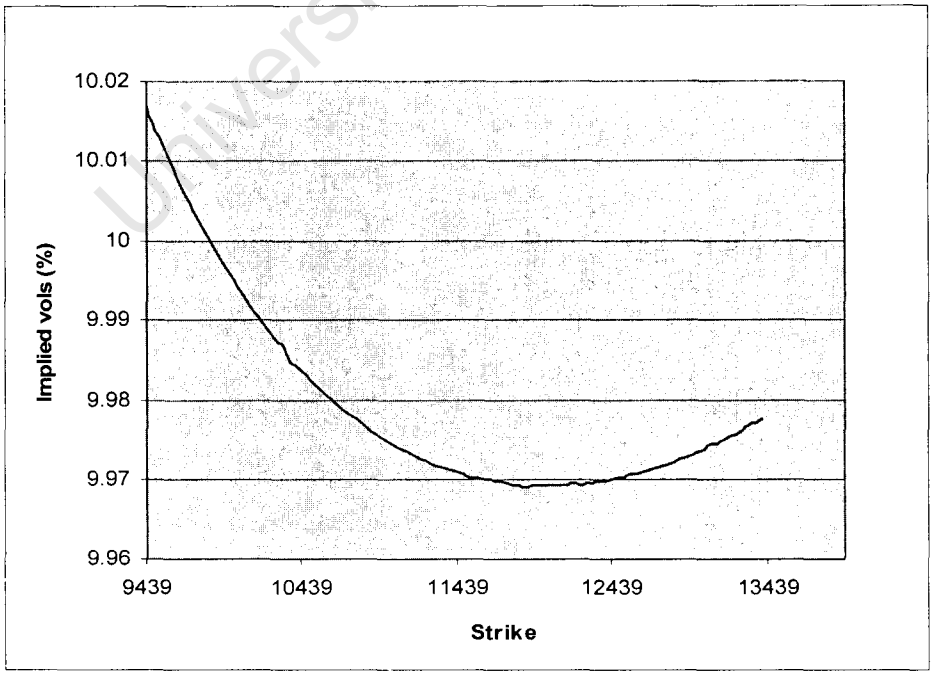


Figure 5.3: The volatility skew produced by NM (2) AGARCH model

The volatility skew shape observed in Figure 5.3 implies that low strike options are more expensive than high strike ones. It is therefore similar to skews observed in most equity markets.

The resulting normal mixture distribution of the equity index returns is shown in Figure 5.4 below along with its constituent densities.

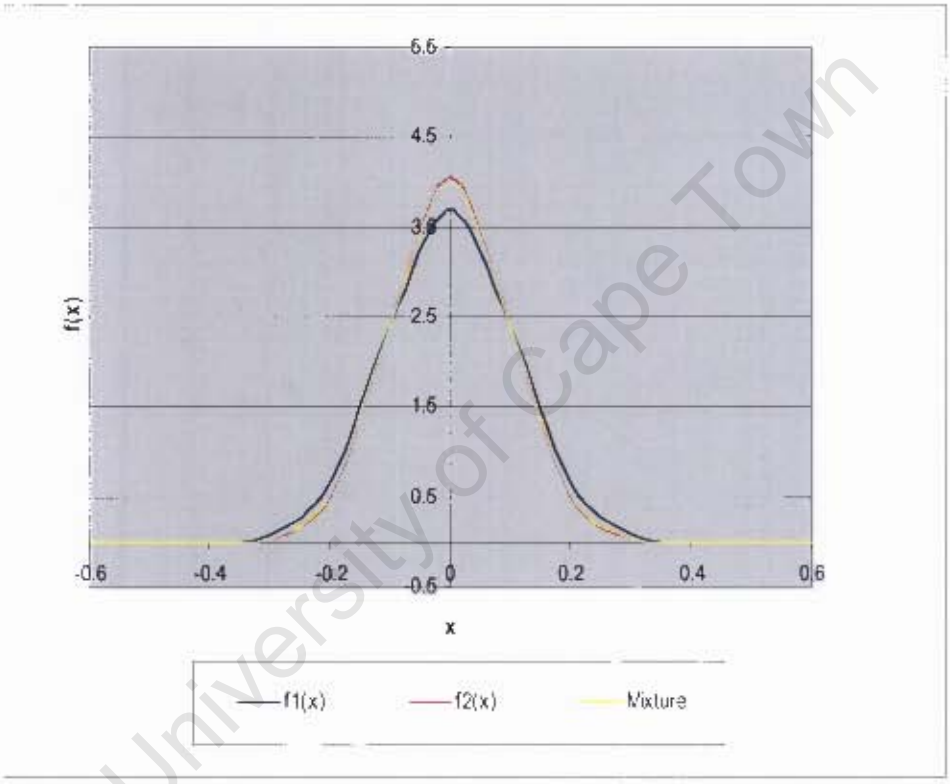


Figure 5.4 : The resulting normal mixture distribution of the index returns

As expected, the normal mixture density in Figure 5.4 is not skewed because it is a mixture of two normal distributions each with zero means. The excess kurtosis of the mixture is calculated to be 0.012. This implies that the mixture distribution has slightly fatter tails than a normal density with the same drift.

In order to evaluate the sensitivity of the NM (2) AGARCH model to changes in its parameter values, each of its parameters was varied individually while the others were kept constant. Each parameter was varied by 1 basis point. This

analysis was used to observe the change in the volatility skew range for slight variations in the NM (2) AGARCH model parameters. Table 5.5 presents the change in the range of the implied volatility skew range (in basis points) as each parameter was varied.

Table 5.5: Variation of skew range as parameters are varied by 1 basis point

Parameters	**Change in range of skew (bps)**
ω_1	0.48
α_1	0.72
β_1	3.74
λ_1	1.32
p	0.03
ω_2	0.5
α_2	0.26
β_2	4.05
λ_2	0.32

Table 5.5 shows that negligible variation in the range of the skew was observed as most parameter values were altered. The table also illustrates that the greatest variation was produced by altering only three of the parameters i.e. β_1, β_2 and λ_1. These observations nevertheless illustrate that the NM (2) AGARCH model is not very sensitive to small changes in its parameter values.

5.2 *Calibration from current option prices*

Call option prices on the JSE TOPI40 Index were used to calibrate both the local volatility mixture and jump diffusion models. The volatility surface for the JSE TOPI40 Index option as observed on the 1 March 2005 is shown in Figure 5.5.

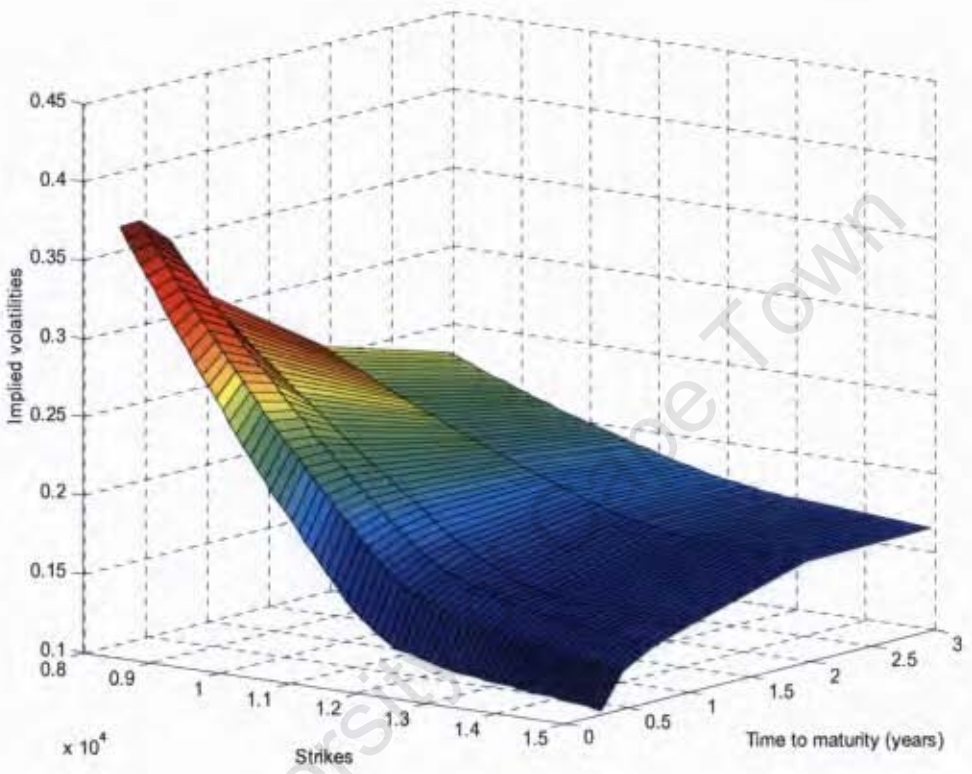


Figure 5.5: The volatility surface for the JSE TOPI40 Index option as observed on the 1 March 2005

Figure 5.5 illustrates that the volatility skew is very steep for short maturity options and becomes flatter as the time to expiry increases. A cross section of the volatility surface was taken i.e. the volatility skew at a maturity of 0.5 years. Each model was then calibrated to this skew. The results are presented in Sections 5.2.1 and 5.2.2. Calibration errors produced by the models are then compared in Section 5.2.3.

5.2.1 Local volatility mixture model results

Examples of the fitting quality of the three models are illustrated below. The inputs shown in Table 5.6 were used to calibrate the local volatility mixture models to market data.

Table 5.6 : Inputs used to calibrate the local volatility mixture models to market data.

Today's date	2005/03/01
Option Expiry Date	2005/09/15
Today's index level	12210
Dividend yield (%)	0
Risk Free Rate (%)	7.5
No of mixtures	3
Call / Put (1/-1)	1

Table 5.6 above shows that the index was expected to pay no dividends over the life of its options. The number of mixtures was set to 3, implying that the index density function was a mixture of 3 lognormal densities.

Using the inputs shown in Table 5.6, call option prices were computed for various strike prices. The local volatility model parameters were found by minimising the squared errors between the market and model prices. The admissible parameter values for the basic lognormal mixture model are displayed in Table 5.7.

Table 5.7: Admissible lognormal mixture model parameter values

i	λ_i	η_i
1	0.0702955	0.6034137
2	0.2326565	0.0978772
3	0.6970479	0.1005125

The resulting calibrated implied volatility skew is displayed in Figure 5.6 where it is compared to the market skew.

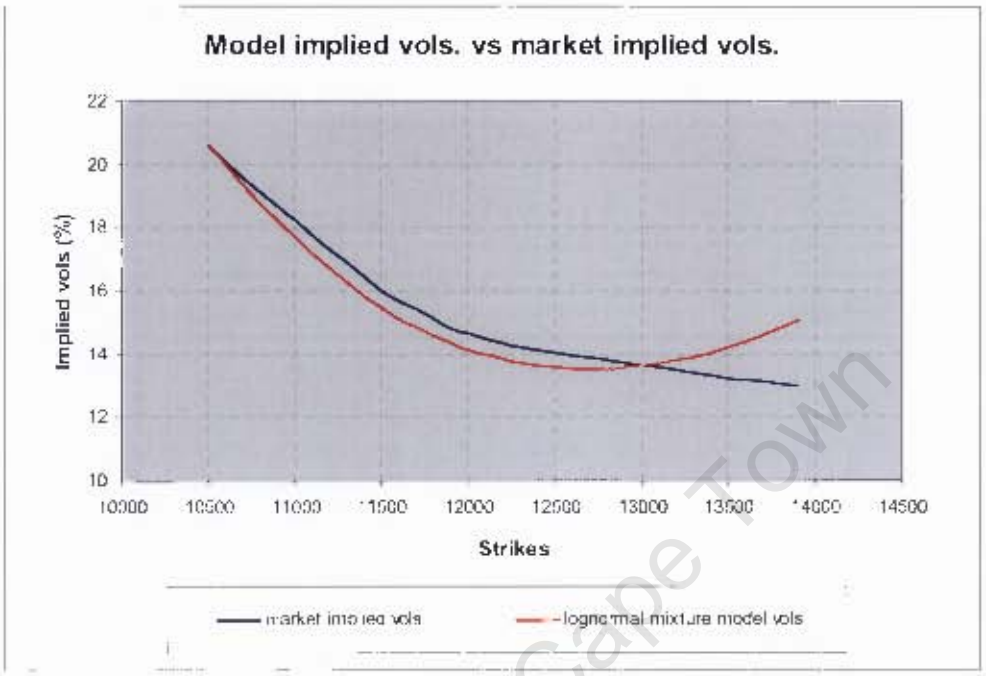


Figure 5.6: Calibrated lognormal mixture model volatilities vs. market implied volatilities

Figure 5.6 illustrates that the lognormal mixture model over prices high strike options relative to the market. The model, however slightly under prices options with strikes less than 13000. The inputs shown in Table 5.6 were also used to calibrate the shifted mixture model to market data. The admissible parameter values for this model are displayed in Table 5.8.

Table 5.8 : Admissible shifted mixture model parameter values

α	-5.40	
i	λ_i	η_i
1	0.9802693	0.0174039
2	0.0000000	0.0308455
3	0.0197306	0.2145585

The resulting calibrated volatility skew is displayed in Figure 5.7 where it is compared to the market skew.

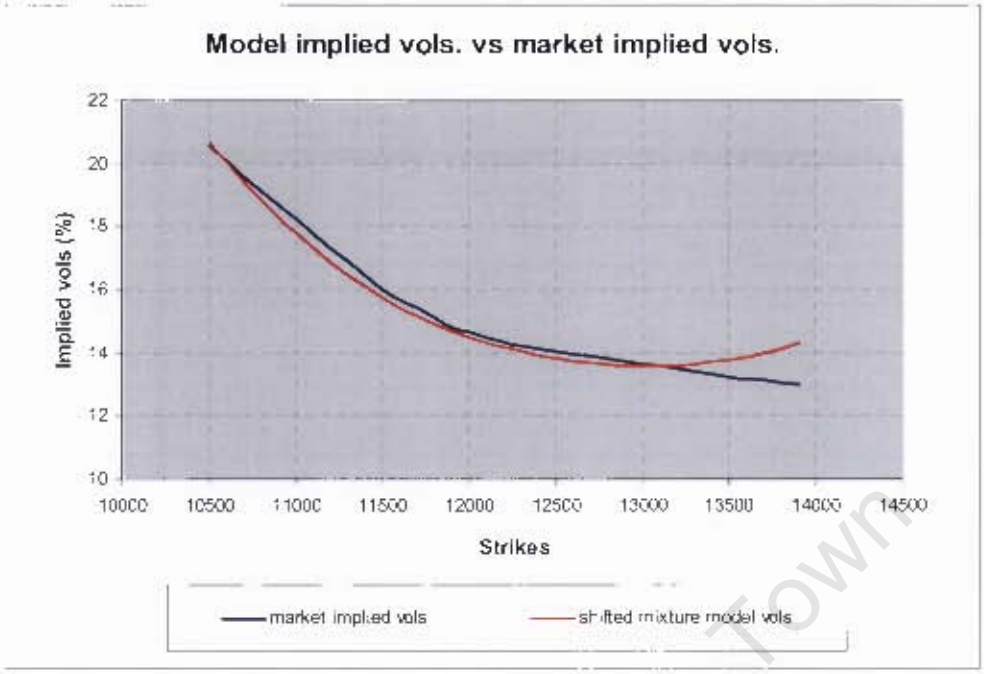


Figure 5.7: Calibrated shifted mixture model volatilities vs. market implied volatilities

Once again, Figure 5.7 illustrates that the shifted mixture model over-prices high strike options relative to the market. For strikes less than 13000, however, the implied volatilities of this model match the market implied volatilities more closely than the basic lognormal mixture model.

After calibrating the lognormal mixture model (with different means), the parameter values displayed in Table 5.9 below were obtained.

Table 5.9: Admissible lognormal mixture model (with different means) parameter values

i	λ_i	η_i	$M_i(T)$
1	0.0000000	0.1000000	-0.0846901
2	0.6499856	0.0757704	0.0526203
3	0.3500143	0.2394345	0.0181358

The resulting model volatility skew is displayed in Figure 5.8.

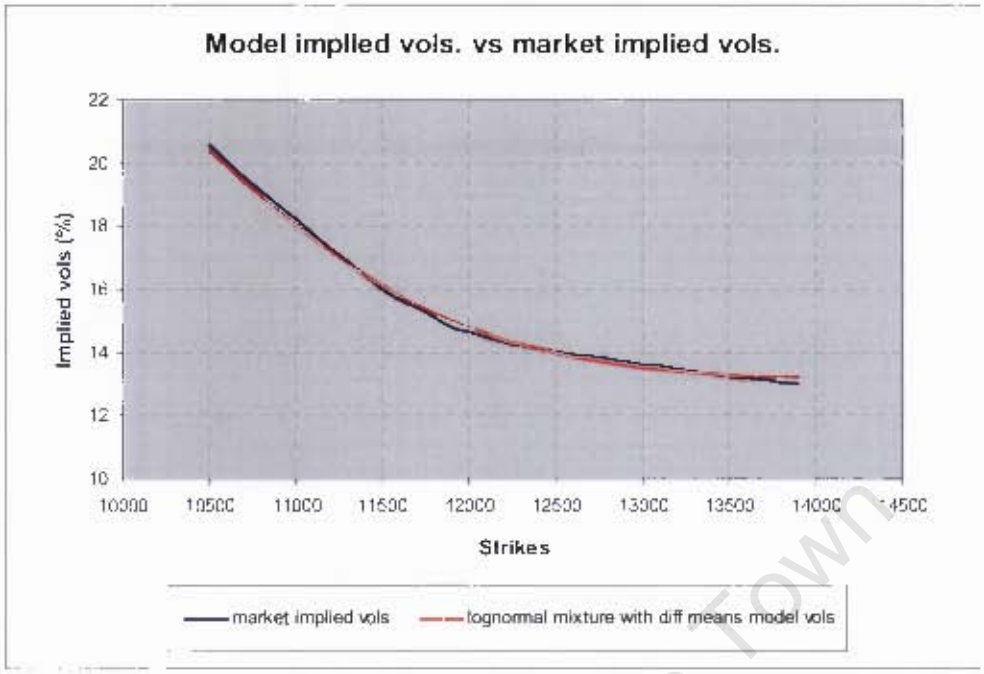


Figure 5.8 : Calibrated lognormal mixture model (with different means) volatilities vs. market implied volatilities

Figure 5.8 shows that the implied volatilities of this model match the market implied volatilities more closely than the previous local volatility models investigated. The resulting S_T densities produced by the three local volatility mixture models are compared with a lognormal distribution in Figure 5.9.

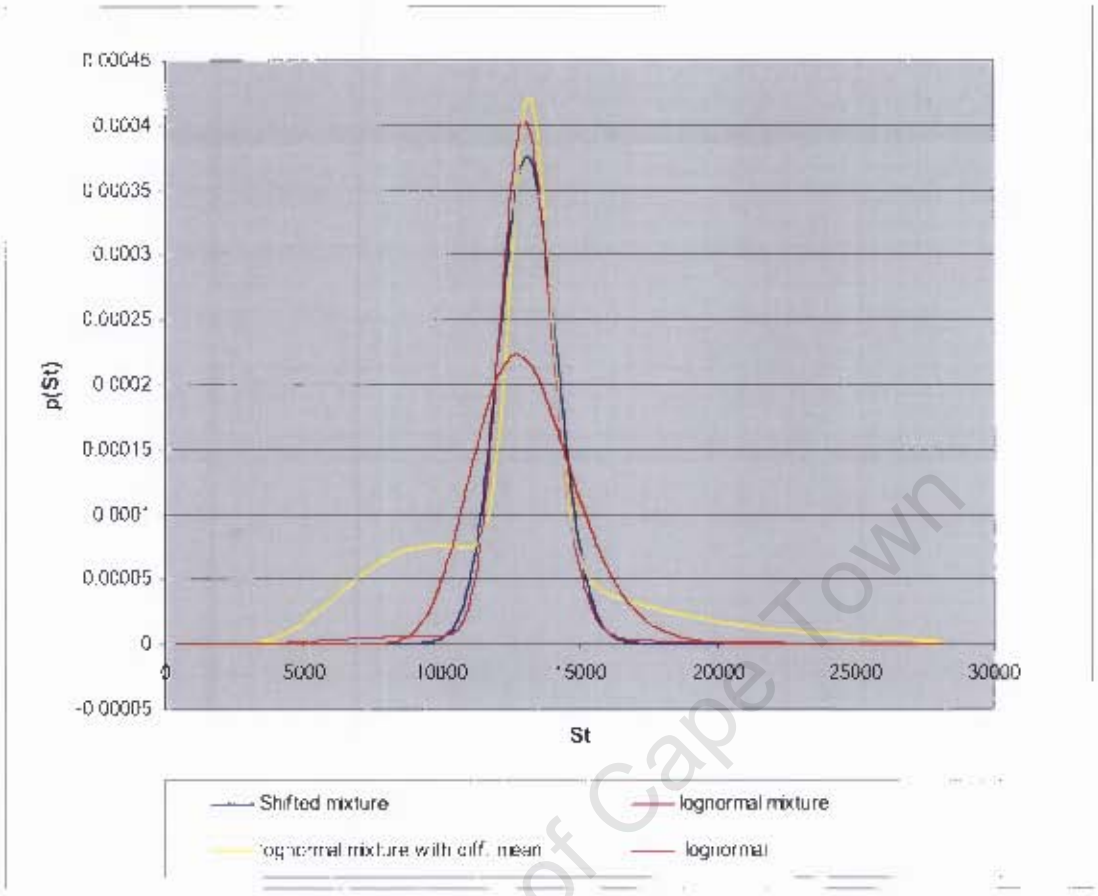


Figure 5.9 : The resulting density p(x) of S_T for the three local volatility models

The densities produced by the local volatility models are more peaked than a lognormal density. They also have fatter left tails and less fat right tails than the lognormal distribution. The skewness and the excess kurtosis of the returns distribution were computed empirically. They were calculated as 0.006383 and 0.281 respectively for the basic lognormal distribution model, 0.007783 and 0.1223 for the shifted mixture model and -1.7429 and 0.331809 for the lognormal mixture model with different means. These results indicate that the returns distribution is more peaked and skewed than a normal distribution.

5.2.2 The double exponential jump diffusion model results

The double exponential jump diffusion model was also calibrated to market data. The fitting quality of the model is illustrated by the example below. The inputs shown in Table 5.10 were used to calibrate the model to market data.

Table 5.10 : Inputs used to calibrate the double exponential jump diffusion model to market data

Today's date	2005/03/01
Option Expiry Date	2005/09/15
Today's index level	12210
Dividend yield (%)	0
Risk Free Rate (%)	7.5
Call / Put (1/-1)	1
No of terms	10

Although the option pricing formula involves an infinite series, Table 5.10 illustrates that the first 10 terms were chosen for this example. This is because empirical work in [30] suggests that a 10 to 15 terms are sufficient for most applications. Admissible parameter values for the model are displayed in Table 5.11.

Table 5.11 : Admissible double exponential jump diffusion model parameter values

η_1	90.15861952
η_2	0.452183953
p	0.940216086
λ	0.393602215
σ	0.123035062

The resulting calibrated skew is displayed in Figure 5.10 where it is compared to the market skew.

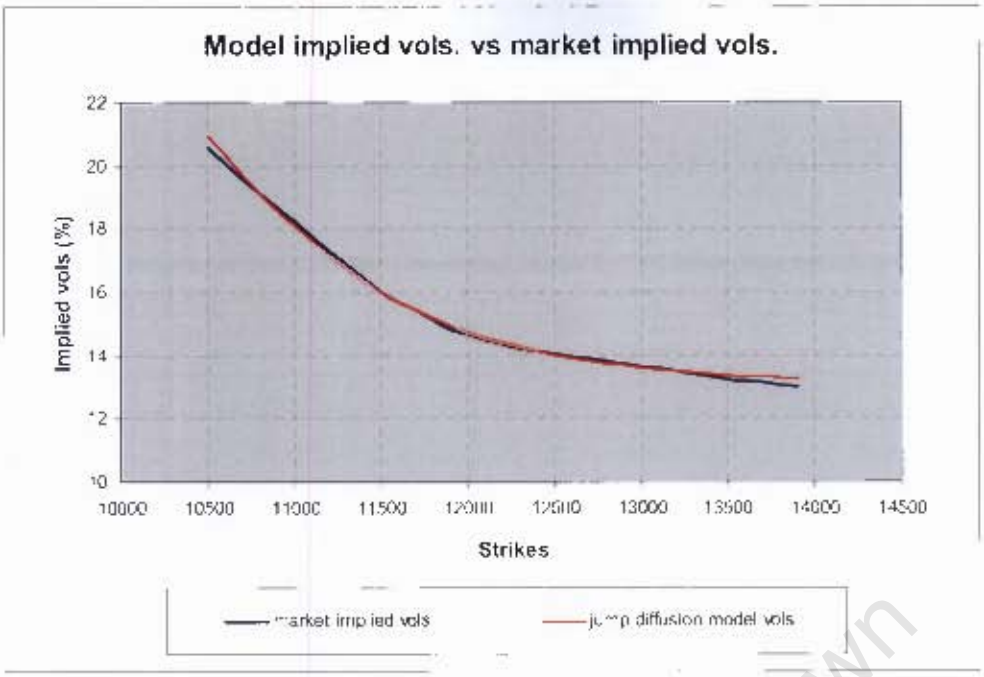


Figure 5.10 : Calibrated jump diffusion volatilities vs. market implied volatilities

Figure 5.9 illustrates the close fit between the jump diffusion skew and the market skew. The jump diffusion model parameter values shown in Table 5.11 were then used to obtain the density function for the approximated returns. This density is plotted in Figure 5.11.

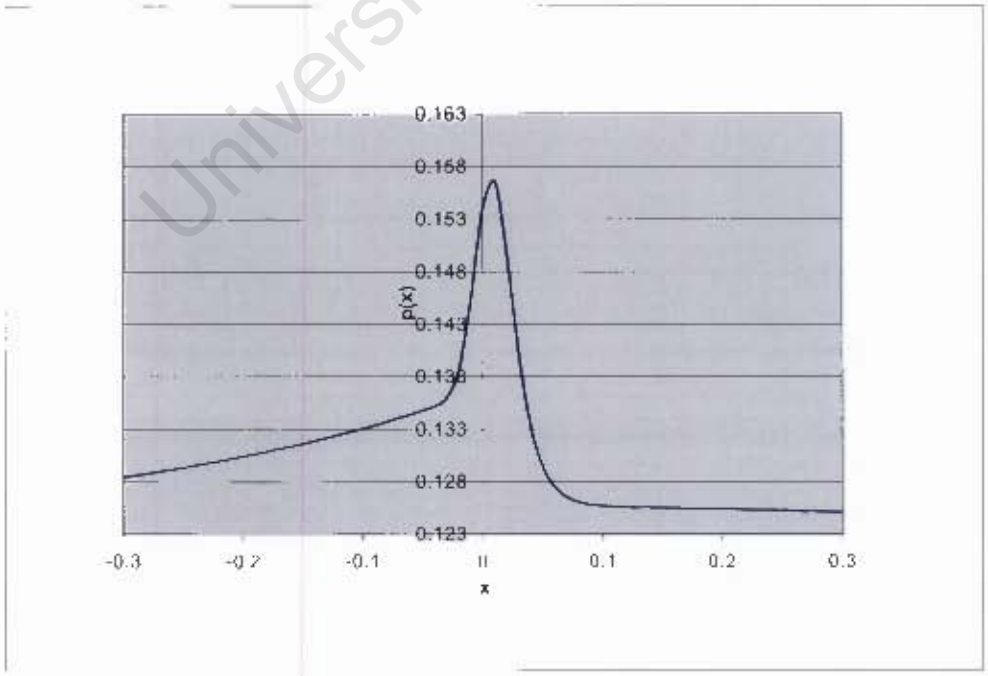


Figure 5.11 : The resulting density function for the approximated returns

The density function in Figure 5.11 is observed to be negatively skewed. Its skewness and excess kurtosis values were calculated as -1.003 and 0.562584 respectively. This density is consistent with the risk neutral density of the underlying asset returns observed in equity markets. Furthermore it is also consistent with the implied volatility skew in Figure 5.10 because low strike options have higher prices than high strike options.

Table 5.12 below compares the skewness and excess kurtosis values of the density functions implied by the local volatility and jump diffusion models.

Table 5.12: Comparing skewness and excess kurtosis

Statistic	Model			
	LM[1]	**SML**	**LMDM**	**DEJD**
Skewness	0.0064	0.0078	-1.7429	-1.003
Excess kurtosis	0.2810	0.1223	0.3311	0.5626

Both the LMDM and DEJD models are negatively skewed. Table 5.12 above also illustrates that both these models have the largest excess kurtosis values.

[1] The models represented by the abbreviations are: LM- basic lognormal mixture model. SML- shifted mixture model of lognormal distributions. LMDM -lognormal mixture with different means. DEJD- double exponential jump diffusion model.

5.2.3 Comparing model calibration errors

Calibration errors for the models were compared using:

I. Different data sets for a specific maturity date
II. Options with different maturities

The errors are expressed in basis points (bps). In this section, the models are abbreviated as follows:

- The basic lognormal mixture model (LM)
- The shifted mixture model of lognormal distributions (SML)
- The lognormal mixture model with different means (LMDM)
- Double exponential jump diffusion model (DEJD)

I. Using different data sets to compare model calibration errors

Three different data sets were used to compare calibration errors. Each set consisted of call option prices (in the form of implied volatilities) on the JSE TOPI40 Index. Only the results for data set 1 are presented in this section. The results for data sets 2 and 3 can be found in Appendix B. Details of the option used as data set 1 are:

Table 5.13: Option details for data set 1

Start Date	1 May 2005
Expiry Date	15 September 2005
Time to expiry (years)	0.3753425
Index Level on start date	11446
Risk free rate (%)	6.95

The calibration errors found using this data set for each model are presented in Table 5.13.

Table 5.14: Calibration error measurements for data set 1

Error measurement	LM	SML	LMDM	DEJD
Max absolute error (bps)	56.14	32.80	31.43	**18.05**
Mean absolute error (bps)	0.47	0.20	0.27	**0.17**
Sum of errors squared (bps)	1.19	0.26	0.36	**0.13**
Root mean square errors (bps)	19.67	9.20	10.84	**6.67**

The results in Table 5.14 show that the double exponential jump diffusion model (DEJD) produces the lowest calibration errors .Figure 5.12 illustrates the volatility skews produced by each model. The market implied skew is also plotted.

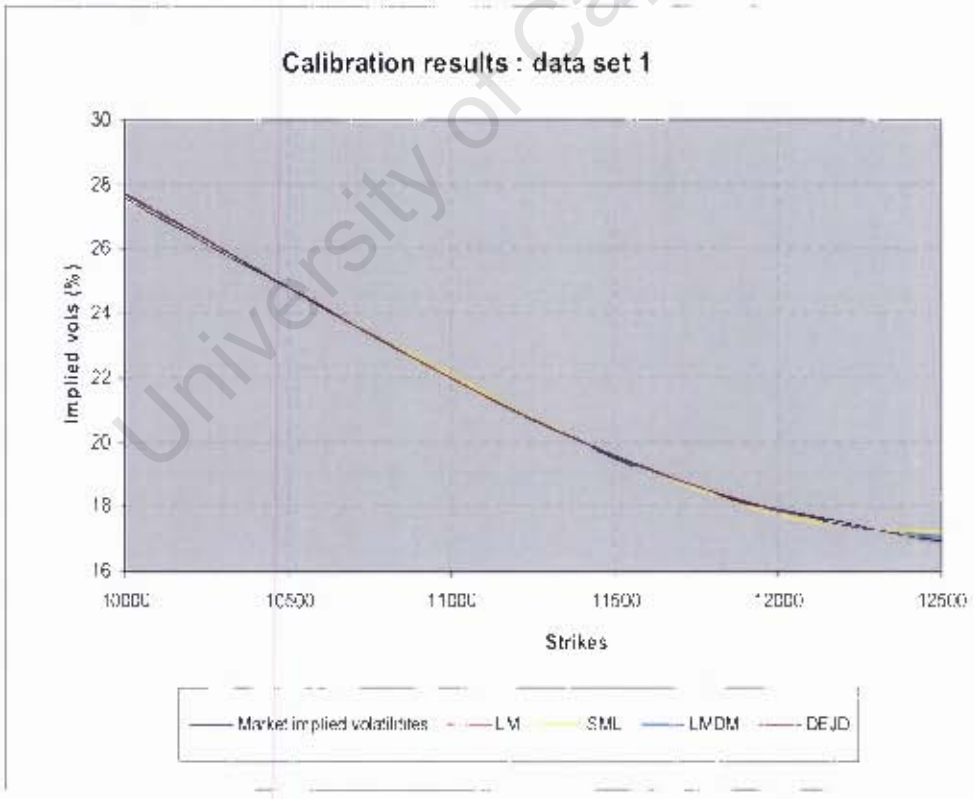


Figure 5.12: Results of model calibration for data set 1

The jump diffusion model (DEJD) gives rise to a more leptokurtotic and negatively skewed asset returns distribution than the local volatility models.

Therefore, as expected, the results from data sets 1-3 illustrate that this model produces the best fit to market implied volatilities. The LMDM and SML models produced comparable results for data sets 1 and 3. For data set 2, however, the LMDM model produced slightly smaller calibration errors. Also, the additional estimated parameters enabled the LMDM and SML models to produce smaller calibration errors than the LM model.

II. Using different option maturities to compare model calibration errors

The local volatility and jump diffusion models were calibrated to market implied data of options with different maturity dates. The inputs to each model were:

Start Date	1 April 2005
Index Level on start date	12210
Risk free rate (%)	7.59

Three ranges of time to maturity were distinguished:

- Short maturity (below 0.5 years)
- Medium maturity (between 0.5 and 1 year)
- Long maturity (above 1 year)

Root mean square errors (RMSE) and maximum absolute errors (ABSERR) across these maturities are displayed in Figures 5.13 and 5.14 respectively. Although not displayed, similar trends were observed for the sum of square errors (SSE) and absolute mean errors (ABSMEAN). The calculated errors using all four measures can be found in Appendix C.

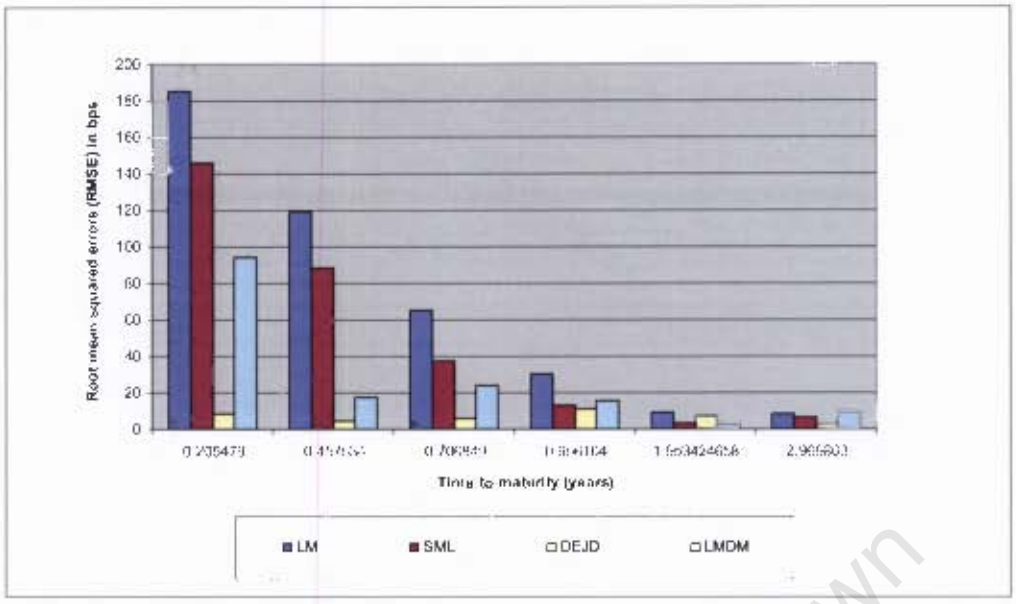


Figure 5.13: Comparing root mean squared errors produced for options with different maturities

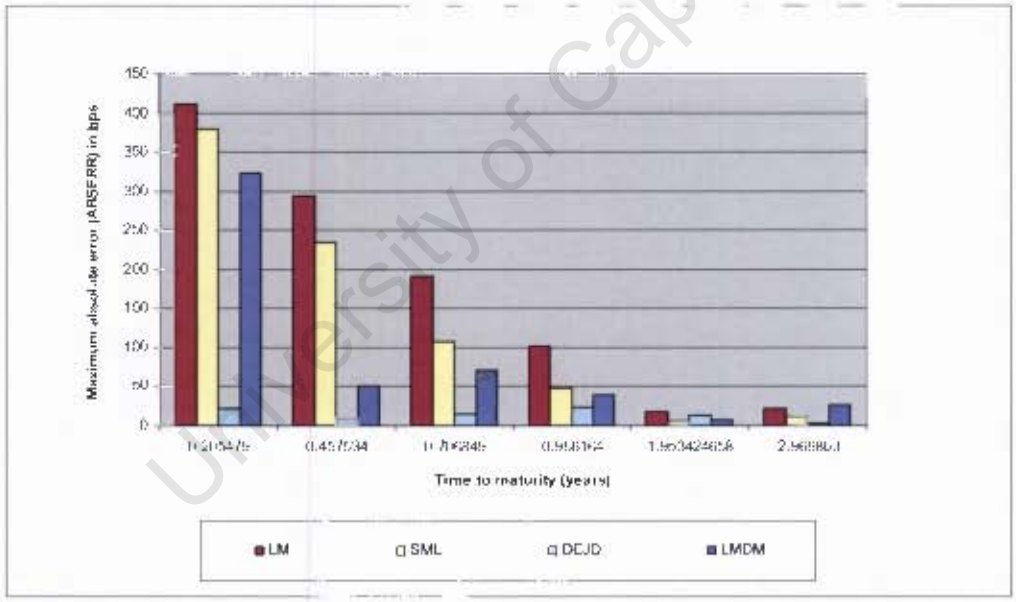


Figure 5.14: Comparing maximum absolute errors produced for options with different maturities

For different maturities, the results in this section show that the DEJD model produced smaller RMSE and ABSERR errors than all three local volatility models. The LM model, on the other hand, produced the largest errors for most maturities. RMSE and ABSERR errors for the LMDM model were smaller than those of the SML model for short and medium maturity options. For long maturity options, the LMDM model produced slightly larger errors.

It was also observed that the DEJD model calibration errors remained relatively constant across all maturities. RMSE for the model were found to remain below 20 bps while ABSERR stayed under 30 bps for all maturities investigated in this thesis. Calibration errors for the local volatility models, on the other hand, decreased drastically as option maturity was increased. Although it appears that calibration errors decrease as the time to maturity of the option increases, the results might not be entirely accurate for long maturity options. This is because long maturity options on the JSE TOP40 index are not as liquid as short and medium maturity options

Each model was also calibrated using a different data set. RMSE and ABSERR errors for different maturities were computed using this data set. The results are presented in appendix D. In general, the DEJD model was found to perform better than the local volatility models. The LM model produced the largest calibration errors. For short and medium maturity options, the LMDM model produced smaller calibration errors than the SML model. The opposite result was obtained for long maturity options.

Once again, for the local volatility models, calibration errors were observed to decrease drastically as the option maturity increased. RMSE for the DEJD model remained under 20 bps while ABSERR stayed under 30 bps for all maturities.

6 Conclusions

The main aim of the empirical work done in this dissertation was to reproduce the volatility skew observed in the South African equity markets. In order to achieve this objective, four volatility models were investigated.

The parameters of two of these models i.e. the Edgeworth and NM (2) AGARCH models were estimated from historical prices of the JSE TOPI40 index. These models are most useful to asset managers as they enable them to obtain a theoretical model of the volatility skew prior to consulting a broker. Additionally, by altering their assumptions regarding the distribution of the underlying, different volatility skews can be obtained. The aim of parameter estimation via historical data was to observe the volatility skew produced and the distribution of the underlying index returns.

A sensitivity analysis was used to determine the sensitivity of each model to variations in the input parameters, in the vicinity of their best estimates. This was the only means used to evaluate the performance of these models.

The remaining models i.e. the local volatility mixture and double exponential jump diffusion models were calibrated using the current prices of liquid call options on the JSE TOPI40 Index. With these models, users such as banks or corporations can evaluate how the market is pricing risk relative to the model. Thus offering a useful tool to determine whether to buy or sell an option.

Four objective measures were used to evaluate and compare the fitting quality of each model. The distribution of the underling index returns was also observed.

Volatility skew modelling from historical prices

Based on the findings the following conclusions can be drawn:

- Both the Edgeworth and NM (2) AGARCH models can produce volatility skew shapes that closely resemble the skews observed in South African equity markets.

- The Edgeworth model is not sufficiently robust. The volatility skew produced by this model is very sensitive to slight changes in the model parameters values.

- The NM (2) AGARCH model, in contrast, is not sensitive to small changes in its parameter values. However, the model cannot be used to back out a risk neutral density and therefore the resulting skew cannot actually be used to price options.

Volatility skew modelling from current option prices

From the results of the calibration from current option prices the following conclusions can be drawn:

- Both the local volatility mixture and double exponential jump diffusion models can be successfully calibrated to market data.

- The densities of the underlying index returns implied from the local volatility mixture and jump diffusion models are consistent with those observed in equity markets. Not only are all the densities more peaked than a normal distribution but the densities produced by the lognormal mixture model (with different means) and the double exponential jump diffusion model are also negatively skewed.

- For different data sets, the jump diffusion model produces the best fit to market implied volatilities. The basic lognormal mixture model, on the other hand, produces the worst fit.

- For different maturities, the jump diffusion model produces smaller calibration errors than the three local volatility mixture models investigated. The largest errors are observed when the basic lognormal mixture model is used. Root mean squared and maximum absolute errors for the lognormal mixture model (with different means) are smaller than those of the shifted mixture model for short and medium maturity options. For long maturity options, the lognormal mixture model (with different means) produces slightly larger calibration errors.

  Furthermore, calibration errors for the local volatility models decreased drastically as the option maturity was increased. Jump diffusion model calibration errors, on the other hand, remained relatively constant across the maturities investigated.

- Maximum absolute errors for the jump diffusion model are found to be below 30 bps across all maturities considered. This makes this model ideal for practical use as it does not need to be recalibrated for different maturities.

- However, calibration errors for the local volatility mixture models vary significantly with time to maturity. As a result, if these models were to be used in practice, it would be essential to recalibrate them for different maturities.

References

1. M. Rubinstein, 'Edgeworth Binomial Trees'. *Journal of derivatives*, vol.5 no.3, pp. 20- 27, 1998.
2. R. Jarrow and A. Rudd. 'Approximate option valuation for arbitrary stochastic processes.' *Journal of Financial Economics*, vol. 10, pp.347-369, 1982
3. M. Jackson & M. Staunton. *Advanced modelling in finance* using Excel and VBA. New York: John Wiley & Sons, 2001 (pp212-220).
4. J. Hull. *Options, Futures and Other Derivatives*. 5th ed. New Jersey: Prentice Hall, 2003 (pp 330-341).
5. C. Alexander. Taming the skew. ISMA Centre Discussion Papers in Finance. The School for Financial Markets. The University of Reading. 2001.
6. J. Hull. *Options, Futures and Other Derivatives*. 4th ed. New Jersey: Prentice Hall, 2003 pp 204.
7. D. Madan and F.Milne. 'Contingent claims valued and hedged by pricing and investing in a basis'. *Mathematical Finance* vol. 4, no.3, pp.223-245, 1994.
8. C.J. Corrado and T. Su. 'Implied volatility skews and stock index skewness and kurtosis implied by S&P 500 Index option prices'. *Journal of Derivatives*, vol. 4, no. 4, pp.8-19, 1997.
9. E Derman. 'Regimes of volatility'. *RISK,* vol.4, pp. 55-59, 1999
10. E.Derman and I Kani. 'The volatility smile and its implied tree'. Quantitative Strategies Research, Goldman Sachs, 1994.
11. B. Dupire. 'Pricing with a smile'. *RISK*, vol. 7, pp.18-20, 1994.
12. J. Gatheral. 'Lecture 1: Stochastic volatility and local volatility'. Working paper, Courant Institute of Mathematical Sciences. 2001a
13. J. Gatheral. 'Lecture 2: Fitting the volatility skew'. Working paper, Courant Institute of Mathematical Sciences. 2001b
14. M. Rubinstein. 'Implied Binomial Trees'. *The Journal of Finance*, vol. 49 no.3 pp 771-818, 1994.
15. R. Merton. 'The theory of rational option pricing'. *Bell Journal of Economics and Management Science*, vol. 4 , no. 1, pp. 141-183,1973
16. J. Hull and A. White. 'The Pricing of options on assets with stochastic volatilities'. *The Journal of Finance* vol. 42 no.2 , pp. 281-300, 1987

17. R. Merton. 'Option pricing when the underlying stock returns are discontinuous'. *Journal of Financial Economics, vol.* 3, no. 1, pp.125-144, 1976.
18. B. Dupire. 'Pricing and hedging with smiles'. *RISK*, vol. 7, pp. 18-20, 1993.
19. D. Brigo, F. Mercurio & G. Sartorelli. 'Alternative asset price dynamics and volatility smile', *Quantitative Finance,* vol. 3, pp. 173-183, 2003.
20. D. Brigo & F. Mercurio. 'Skew before it's too late'. *RISK*, vol. 13, no. 9, pp 123-126, 2000.
21. D. Brigo, F. Mercurio & F. Rapisarda. 'Lognormal Mixture Dynamics and Calibration to Market Volatility Smiles'. *International Journal of Theoretical and Applied Finance*, vol. 5, no. 4, pp. 427-446, 2002.
22. D. Brigo, F. Mercurio & G. Sartorelli. 'Alternative asset price dynamics and volatility smile'. *Working paper*. BANCA IMI. [Online]. Available: http://www.damianobrigo.it/ [2007, June 16]
23. D. Brigo & F. Mercurio. 'Displaced and mixture diffusions for analytically tractable smile models'. Mathematical Finance-Bachelier Congress 2000 (Springer Finance) ed German, Madan,Pliska Vorst (Berlin: Springer)pp 151-174
24. C. Alexander. *Market Models: A Guide to Financial Data Analysis.* New York: John Wiley & Sons. 2001.(Ch6 pgs154-171 pg68)
25. S. Aboura. *A Survey on Implied Theories: The Volatility Models.* [Online]. Available: http://papers.ssrn.com/sol3/cf_dev/AbsByAuth.cfm?per_id=101097 2002. [2007, June 4].
26. C. Alexander. 'Local Volatility'. Presented at the University of Rome .Correlation, Cointegration and Calibration Conference 3-5 June 20002.
27. S.L. Heston. 'A Closed-Form Solution for Options with Stochastic Volatility Applications to Bond and Currency Options'. *Review of Financial Studies*, vol. 6, no.2, pp. 327-43, 1993.
28. C. Alexander, G. Brintalos & L. Nogueira. Short and Long Term Smile Effects: The Binomial Normal Mixture Diffusion Model. ISMA Centre Discussion Papers in Finance. The School for Financial Markets. The University of Reading. 2003.

29. J. Rice. *Mathematical Statistics and Data Analysis*. 2nd Ed. California: Wadsworth Publishing Company. 1995. (pp. 43-44)
30. S.G Kou. 'A jump diffusion model for option pricing'. *Management Science*, vol. 48, no.8, pp. 106-1101, 2002.
31. Z. Zhu & F. B. Hanson. 'A Monte Carlo Option-Pricing Algorithm for Log-Uniform Jump-Diffusion Model'. To appear in *Proceedings of Joint 44th. IEEE Conference on Decision and Control and European Control Conference* pp. 1-6. Seville, Spain, 12-15 December 2005.
32. M.I. Abramowitz & A. Stegun, eds. *Handbook of Mathematical Function with formulas, graphs and mathematical tables*. 10th ed. New York: Dover. 1972.
33. A. Sepp and I. Skachkov. 'Option Pricing with Jumps'. *Wilmott Magazine,* vol. , no. pp.50-58, 2003
34. S.G. Kou. 'A jump diffusion model with three properties: leptokurtic feature, volatility smile and analytical tractability'. Contributed to the World Conference Econometric Society, Seattle, August 11-16, 2000.
35. Z. Zhu & F. B. Hanson. 'Comparison of market parameters for jump diffusion distributions using multinomial maximum likelihood estimation'. *Proceedings of 43rd IEEE Conference on Decision and Control,* pp. 3919-3924. Bahamas, December 14-17 2004.
36. T. Bollerslev. 'Generalized Autoregressive Conditional Heteroskedasticity'. *Journal of Econometrics*, vol. 31, no. 3, pp.309-328, 1986.
37. J.C. Duan. 'The GARCH Option Pricing Model'. *Mathematical Finance*, vol. 5, no.1, pp.13-32, 1995.
38. J.C. Duan. 'Conditionally Fat tailed and the volatility smile in options'. Working paper. Rotman School of Management, University of Toronto. 1999.
39. S.L. Heston and S. Nandi. 'A Closed Form GARCH Option Valuation Model'. *Review of Financial Studies*, vol. 13, pp. 585-625, 2000.
40. C. Alexander & E. Lazar. The Equity Index Skew and Asymmetric Normal mixture GARCH. ISMA Centre Discussion Papers in Finance. The School for Financial Markets. The University of Reading. 2004.
41. C. Alexander & E. Lazar. 'Normal mixture GARCH (1, 1): Applications to exchange rate modelling'. ISMA Centre Discussion Papers in Finance. The School for Financial Markets. The University of Reading.2004.

42. M. Haas, S. Mittnik & M.S. Paolella. 'Mixed Normal Conditional Heteroskedasticity', *Journal of Financial Econometrics, vol.* 2, no.2, pp.211-250, 2004.

43. P. Wilmott. *Paul Wilmott on Quantitative Finance.* New York: John Wiley & Sons, 2000 (pg158)

44. R. I. Becker. 'Financial Mathematics in Excel' .Lecture notes for Numerical Modelling 1.Departement of Mathematics and Applied Mathematics. University of Cape Town. 2004.

45. F. Black & M. Scholes. *'The pricing of options and corporate liabilities'*, Journal of Political Economics, vol. 81, pp. 637-653.

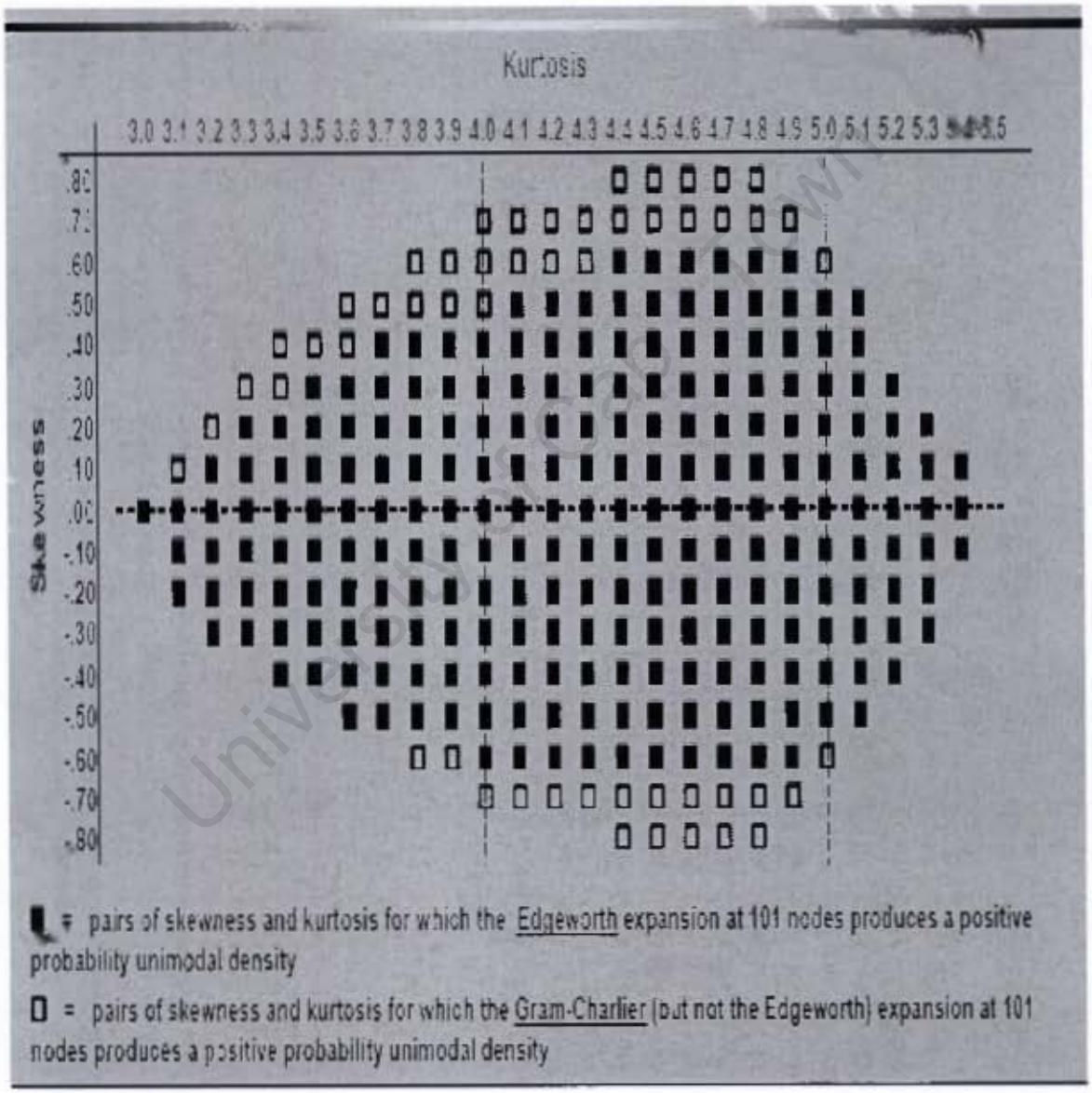
46. S. Natenberg. *Option volatility and pricing: Advanced trading strategies and techniques.*2nd ed. New York: *McGraw-Hill, 1994.*

47. E. Derman. *My Life as Quant. Reflections on physics and finance*. New Jersey: John Wiley & Sons, 2004.

48. N.Johnson et al. *Continuous univariate distributions.* Vol. 1, 2nd Ed. New York: John Wiley & Sons, 1994.

Appendix A

Table A1: The (ξ, κ) pairs which allow for positive unimodal distributions



Appendix B

Using different data sets to compare model calibration errors

Data set 2:

Start Date	1February 2005
Expiry Date	15 December 2005
Time to expiry (years)	0.86849
Index Level on start date	11700
Risk free rate (%)	7.5

Table B1: Error measurements for data set 2

Error measurement	LM	SML	LMDM	DEJD
Max absolute error (bps)	69.79	45.06	35.43	12.18
Mean absolute error (bps)	13.83	8.79	5.97	2.88
Sum of errors squared (bps)	1.26	0.52	0.34	0.04
Root mean square errors (bps)	20.54	13.20	7.69	3.83

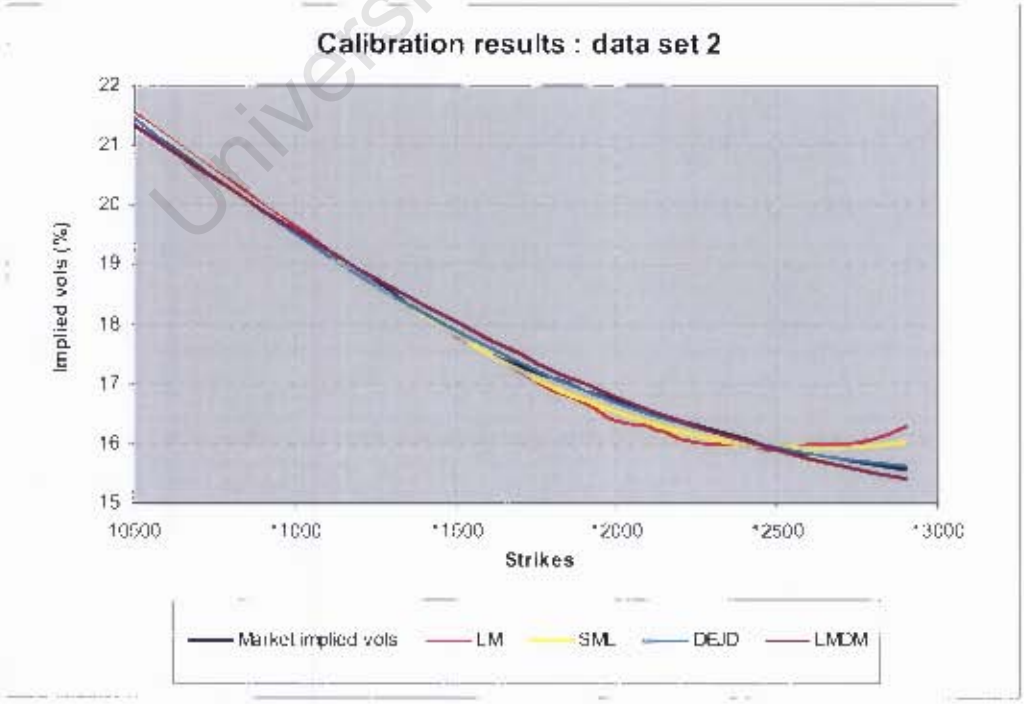


Figure B1: Results of model calibration for data set 2

Data set 3:

Start Date	1 April 2005
Expiry Date	20 March 2008
Time to expiry (yrs)	2.969863
Index Level on start date	12210
Risk free rate (%)	7.59

Table B2: Error measurements for data set 3

Error measurement	LM	SML	LMDM	DEJD
Max absolute error (bps)	11.60	21.15	26.33	3.47
Mean absolute error (bps)	5.89	7.14	7.35	2.10
Sum of errors squared (bps)	0.14	0.23	0.30	0.17
Root mean square errors (bps)	6.54	8.51	9.64	2.30

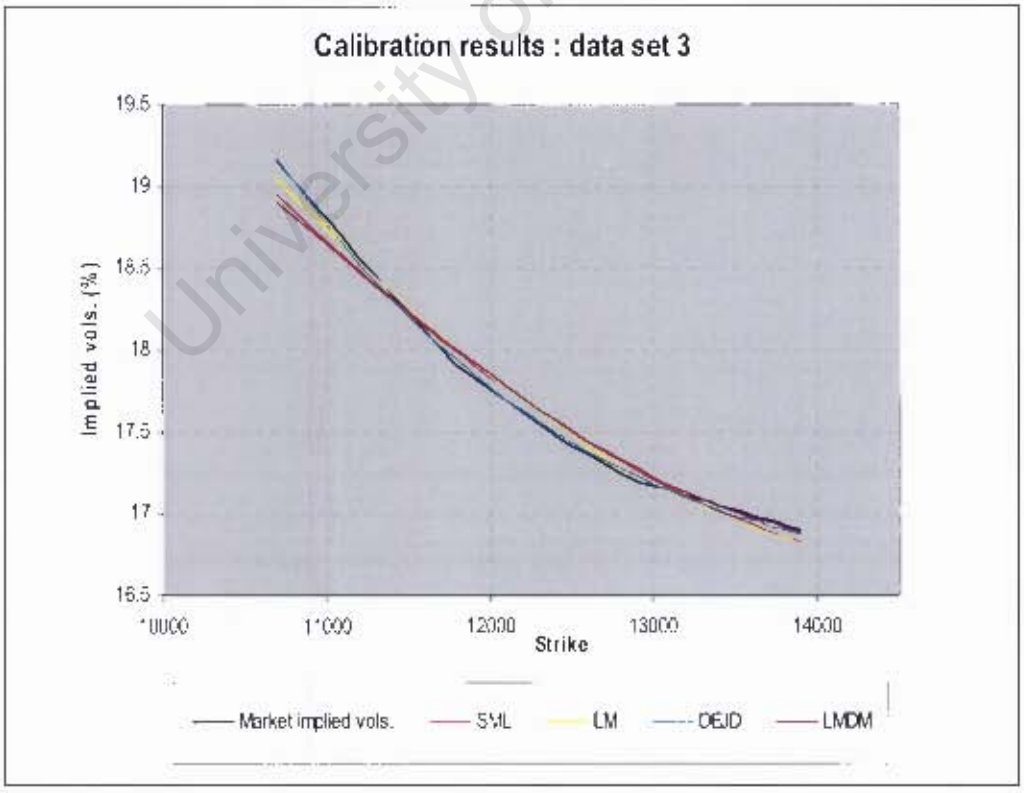


Figure B2: Results of model calibration for data set 3

Appendix C

Calculated calibration errors for the:

- The lognormal mixture model (LM)
- The shifted mixture model of lognormal distributions (SML)
- The lognormal mixture model with different means (LMDM)
- Double exponential jump diffusion model (DEJD)

The calculated errors are:

- Maximum absolute errors (ABSERR)
- Mean absolute errors (ABSMEAN)
- Sum of squared errors (SSE)
- Root mean squared errors (RMSE)

1. Calibration errors produced by short maturity options

Table C1: Calibration errors for short maturity options

Error Measurement (bps)	Time to maturity (years)							
	0.20				0.45			
	LM	**SML**	**LMDM**	**DEJD**	**LM**	**SML**	**LMDM**	**DEJD**
ABSERR	411.08	378.49	322.09	21.24	293.83	234.61	50.18	10.57
ABSMEAN	150.99	113.16	61.17	6.30	101.91	72.99	12.85	3.78
SSE	113.17	70.39	29.36	0.21	46.86	25.85	1.00	0.06
RMSE	185.19	146.05	94.32	8.12	119.16	88.51	17.44	4.47

2. Calibration errors produced by medium maturity options

Table C2: Calibration errors for medium maturity options

Error Measurement (bps)	Time to maturity (years)							
	0.70				0.95			
	LM	**SML**	**LMDM**	**DEJD**	**LM**	**SML**	**LMDM**	**DEJD**
ABSERR	190.66	106.94	70.51	15.17	101.47	47.85	38.51	23.45
ABSMEAN	48.82	30.31	17.87	4.73	20.16	8.77	13.10	9.84
SSE	14.07	4.67	1.92	0.10	3.01	0.56	0.79	0.40
RMSE	65.30	37.65	24.17	5.72	30.23	13.02	15.49	11.09

3. Calibration errors produced by long maturity options

Table C3: Calibration errors for long maturity options

Error Measurement (bps)	Time to maturity (years)							
	1.95				2.96			
	LM	**SML**	**LMDM**	**DEJD**	**LM**	**SML**	**LMDM**	**DEJD**
ABSERR	17.41	5.90	7.79	13.58	11.60	21.15	26.33	3.47
ABSMEAN	8.06	2.49	2.48	6.63	5.89	7.14	7.35	2.10
SSE	0.25	0.20	0.32	0.18	0.14	0.23	0.31	0.17
RMSE	8.86	2.91	3.11	7.39	6.54	8.51	9.64	2.30

Appendix D

Using different option maturities to compare model calibration errors

The following inputs were applied to each model:

Start Date	1 March 2005
Index Level on start date	12202
Risk free rate (%)	7.5

The results for this data set are presented Root mean square errors (RMSE) and maximum absolute errors (ABSERR) for different maturities are displayed in Figures D1 and D2 respectively.

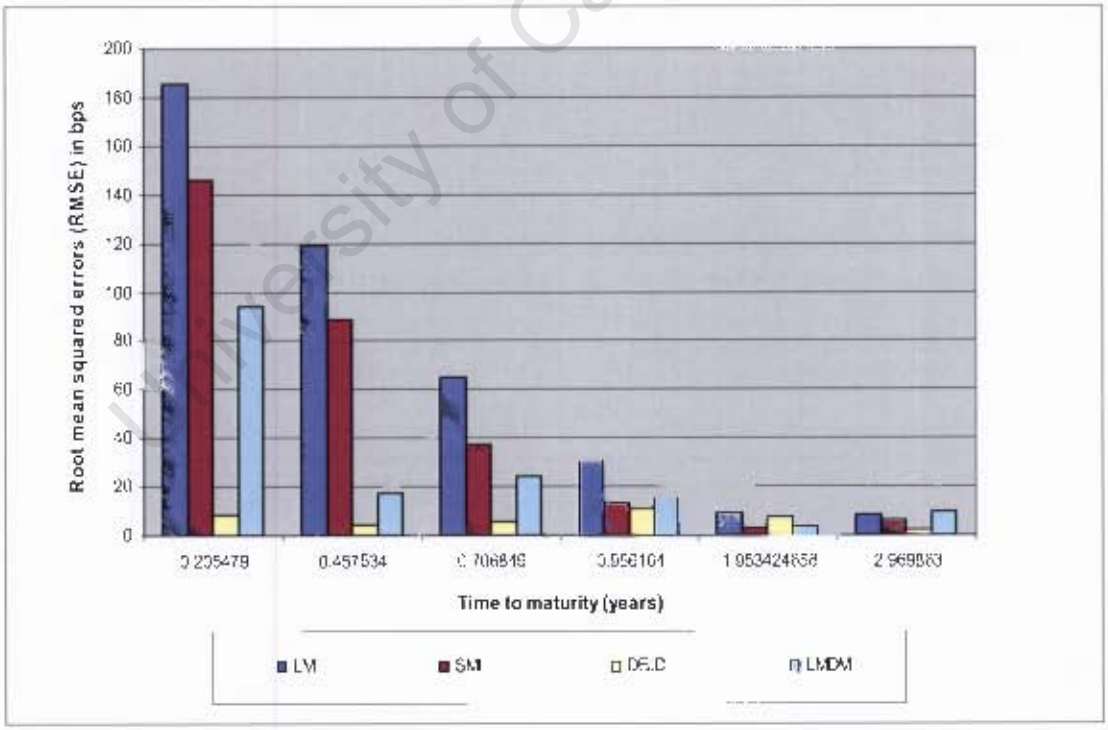


Figure D1: Comparing root mean squared errors produced for options with different maturities

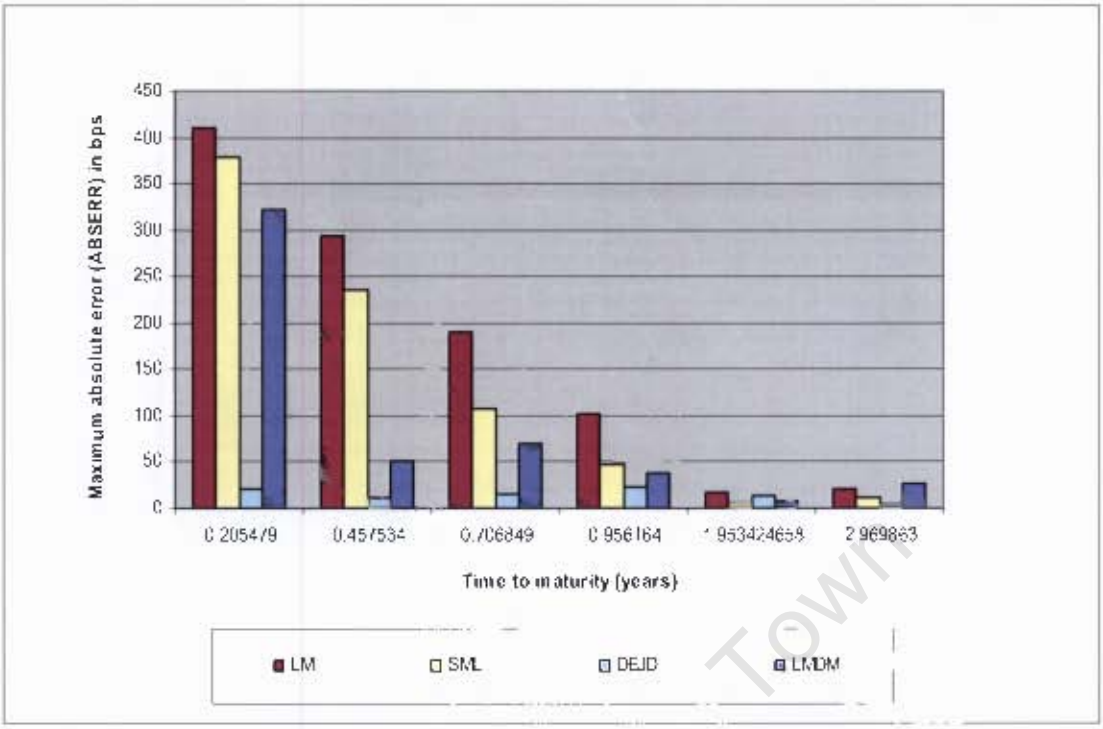


Figure D2: Comparing maximum absolute errors produced for options with different maturities

In general, Figures D1 and D2 illustrate that the double exponential jump diffusion model (DEJD) performed better than the local volatility models. The lognormal mixture model with different means (LMDM) also produced smaller calibration errors than the other two mixture models. The largest errors where observed when the lognormal mixture (LM) model (with same means) was used. Calibration errors were observed to decrease as the option maturity increased.