IS BLACK-SCHOLES THE BEST?
AN ASSESSMENT USING ALTERNATIVE PRICING MODELS

NIKHIL SHENAI

A thesis submitted in fulfilment of the requirements for the degree of MSc Finance and the Diploma of Imperial College London

18th August 2004

SYNOPSIS

This thesis is an assessment of the Black-Scholes option pricing model and some of its alternatives with respect to European calls. Following an examination of Black-Scholes deficiencies, promising alternative models which can correct these defects are drawn from the encompassing time-changed Lévy framework of Carr and Wu. These are Heston's model of Stochastic Volatility, Merton's Jump Diffusion model, the Variance Gamma model of Madan and Seneta, and the Finite Moment Logstable Process of Carr and Wu.

Fast Fourier Transforms and Nonlinear Least Squares are used to calibrate the models quickly, using FTSE 100 call data from the first quarter of 2000. Predicted pricing errors and implied volatility surfaces demonstrate that all the alternative models outperform Black-Scholes. The finite variation of the Variance Gamma model causes it to have the worst performance of the alternative models, while the flexibility of Heston's model enables it to perform the best.

ACKNOWLEDGEMENTS

I would like to thank my family for their support and understanding.

I would like to thank the staff of Imperial, especially my supervisor Dr Aleš Černý, for his help and guidance, and Dr Lina El-Jahel and Dr Dirk Nitzsche, for their comments while Aleš was away.

CONTENTS

NOTATION

BS		**Black-Scholes**
↳	A1-10	Assumptions made in the derivation of the BS formula
	C	Call Price
	d_{1-2}	Arguments for N in BS formula
	dB	A standard Brownian Motion
	GBM	Geometric Brownian Motion
	K	Exercise/Strike Price
	μ	Mean of Share Returns
	p	Put Price
	N	Cumulative Normal Distribution Function
	r	Continuously Compounded Risk-free Interest Rate
	S	Share/Stock Price
	σ	Standard Deviation of Share Returns
	T	Expiration Time
	t	Current Time
	τ	Time to Expiration; $\tau = T - t$
HSV		**Heston's Stochastic Volatility**
↳	dB_1	Standard Brownian Motion in share price process
	dB_2	Standard Brownian Motion in volatility process
	κ_v	Mean-reversion speed
	σ_v	Volatility of volatility
	$\bar{v}$	Long-run variance
	$v(t)$	Stochastic variance
LS		**Finite Moment Logstable Process**
↳	$dL_{\alpha,\beta}$	Levy motion
	dq	Poisson jump process
MJD		**Mixed Jump Diffusion**
↳	$J(\alpha)$	Poisson jump process with intensity α
	k	Jump magnitude with mean θ and variance h^2
VG		**Variance Gamma**
↳	α	Variance of $d\tilde{t}$, controls kurtosis
	$d\tilde{t}$	Subordinated gamma time
	θ	Drift of standard Brownian motion, controls skewness

INTRODUCTION

In 2001, the London International Financial Futures and Options Exchange (LIFFE) traded a record-breaking daily average of €603 billion across 216 million contracts[1]. Rapid growth in the market for derivative products has been fuelled by agents' need to limit their risk in ever more areas, from standard equities to weather, and has resulted in the creation of increasingly sophisticated financial instruments.

Furthermore, the option framework has evolved into a means for analysing any decision with limited downside risk and large upside potential (or vice versa). Merton (1998) shows that the scope for applications is vast, spanning the valuation of corporate liabilities; the pricing of insurance and contract guarantees; tax, legal and human capital analysis; and firms' investment decisions (real options). Of the last application, Pindyck and Dixit (1995, pp. 106) say that the option framework effectively values the flexibility of being able to defer or alter the scale of investment, extending the NPV framework to be able to deal with real life cash-flows more accurately.

Consequently, a clear understanding of how to price options yields not only the ability to guard more precisely against different states of nature, preventing losses to firms as well as individuals investing in pension funds, but also enables more accurate decision-making, promoting greater economic efficiency.

Yet the Black-Scholes option pricing model, the baseline approach to pricing European options, is less of an exact pricing rule, and more of a guideline, as is evident from consistent deviations between Black-Scholes predictions and observed prices. A variety of different models and techniques have arisen which seek to explain these deviations. It is the aim of this thesis to empirically compare the Black-Scholes option pricing model with the most promising of these alternatives.

[1] Based on Liffe.data, a report available at www.Liffe.com

The Literature Survey is comprised of Sections 1-6. Section 1 gives a broad overview of the different approaches involved in option pricing, and introduces the methodology used for investigation. Section 2 describes the Black-Scholes option pricing model, while some of its shortcomings are covered in section 3. Section 4 describes alternative models, the most promising of which are explored further in section 5. Section 6 provides a brief run-through of a framework for evaluating the models, and offers a provisional conclusion and plan of research.

Implementation takes up Sections 7-11. Sections 7 and 8 explain characteristic functions and their option pricing applications in Fourier transforms. Section 9 discusses the methodology used for estimation and testing, and Section 10 briefly discusses the dataset and programs used. The results are presented in Section 11. Section 12 concludes.

1. OPTION PRICING METHODOLOGIES

Many dimensions exist in the option pricing framework. To begin with, Rydberg (1999, pp. 15-16) highlights theoretical discipline, distinguishing between “Mathematical Finance”, which makes use of diffusion processes and stochastic calculus to price derivatives, and “Econometrics”, which involves analysis of time series properties of key variables such as share return variance. This distinction roughly coincides with another boundary, between continuous time and discrete time models. A further difference exists between models with closed-form solutions and those requiring numerical methods. Spectra exist within classes of models too; Mathematical Finance models transition from continuous to pure jump process models, while numerical methods models sort between Monte Carlo simulations and explicit and implicit lattices.

The Black-Scholes model is an example of a continuous time (Mathematical Finance) closed form solution. Other models which have been developed to correct for its perceived deficiencies span the rest of the model subspace (which is still expanding as new approaches are found). According to Malone (2002, pp.22,42) there has not yet been a systematic comparison of existing models across the same dataset. Since the publication of his article, there may have been attempts to conduct such an investigation; indeed, using Levy processes and characteristic functions, Carr and Wu (2002) claim to have derived a framework which encompasses almost all other current models. However, they have not proceeded to empirical testing. Moreover, existing comparisons of option pricing models almost exclusively focus on US data.

It is therefore the aim of this thesis to compare models spanning as much of the model space as possible, but using UK data in order to provide new insights. This in itself is a large remit, so that exploration of other areas, such as the option product space, is unfortunately beyond the scope of the thesis; the focus will be on European share options, for which the model space is most dense.

In order to compare models, a hypothetico-deductive methodology will be implemented, which fits in with the LSE approach to econometric testing as advocated by Hendry (1987). Harvey (1990, pp. 5-6) and Thomas (2000, pp. 361-363) identify desirable features of models following this approach, the most important of which are predictive power, theoretical consistency, data coherence, ability to encompass rival formulations, and parsimony.

To some extent, the literature survey sections will examine models with regard to these features, especially theoretical consistency, ability to encompass rival formulations and parsimony. Having selected diverse models which are promising in these respects, further empirical analysis will be conducted to test for the rest of the features. In particular, models' predictive power will be explored by comparing pricing errors. Data coherence will be checked through tests of parameters.

To start with, the Black-Scholes model and its deficiencies will be outlined.

2. THE BLACK-SCHOLES OPTION PRICING MODEL

In their groundbreaking paper, Black and Scholes (1973, pp.640) made the following explicit assumptions in their derivation:

A1. The share price (*S*) follows a geometric Brownian motion (GBM) process with

constant mean (μ) and volatility (σ), i.e. $\frac{dS}{S} = \mu dt + \sigma dB$

A2. The risk-free rate of interest (*r*) is constant and the same for all maturities.

A3. The option is a European (let call price = *C*).

A4. There are no transaction costs or taxes.

A5. The short selling of securities with the full use of proceeds is permitted.

A6. All securities are perfectly divisible.

A7. There are no dividends or other payoffs during the life of the option.

Implicit in their approach are the following further assumptions:

A8. Security trading is continuous.

A9. There are no riskless arbitrage opportunities.

A10. The stock market is efficient.

The key insight of Black and Scholes was that no-arbitrage pricing can be implemented via a riskless portfolio, created by purchasing one share and selling $\frac{\partial C}{\partial S}$ {= Δ} calls on that share. Imposing the no arbitrage condition (A9), the portfolio's rate of return must equal the constant risk-free rate of interest (A2).

They then used Itô's Lemma to derive the behaviour of the call option price (*dC*) from the share price process (A1). Consolidating algebra, they attained the Black-Scholes partial differential equation:

$$\frac{\partial C}{\partial t} + rS\frac{\partial C}{\partial S} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 C}{\partial S^2} = rC$$

Finally, they imposed the key boundary condition of C = max(S – K, 0) at maturity T (A3) to find the unique closed-form solution:

$$C = SN(d_1) - Ke^{-r\tau}N(d_2)$$

where $$d_1 = \frac{\ln\left(\frac{S}{K}\right) + r\tau}{\sigma\sqrt{\tau}} + \frac{1}{2}\sigma\sqrt{\tau}$$

$$d_2 = d_1 - \sigma\sqrt{\tau}$$

τ = time to expiration

$N(d_i|_{i\,=\,1,2})$ is the cumulative probability function for a standardised normal variable evaluated at d_i

[puts are evaluated by via put-call parity]

3A. VALIDITY OF BLACK-SCHOLES ASSUMPTIONS

Violation of the assumptions presented before can lead to pricing and hedging errors, necessitating extensions to the Black-Scholes (BS) model where it is possible to incorporate a more realistic assumption into the framework, or a completely alternative model where it is not. The assumptions will now be evaluated, in order to determine which models most merit further investigation.

To begin with, the more realistic or less significant assumptions will be covered. A5 holds in real life, though Hull (1997) says that short selling is only permitted when the most recent movement of the share was an increase. A6 does not hold perfectly due to partial incompleteness of markets but is a close approximation to reality so its relaxation will not cause large errors. A7 and A8 are violated respectively by dividends and exchange closures but these can be adequately corrected for by adding them back to share prices and reducing the value of τ used (e.g. from number of days till expiration divided by 365 to number divided by 250).

A10 is the motivation for modelling the share price as a Markov process. Copeland and Weston (1992, pp. 392) show that various studies support weak to semi-strong efficiency in the stockmarket, while Mishkin (2003) further supports this viewpoint with the observation that few investment analysts and mutual funds are able to predict future prices consistently. However, it will be noted that unless stockmarket efficiency holds with certainty, a Joint Hypothesis problem may arise later when testing models; bad performance by a model may not be distinguishable from stockmarket inefficiency because both will cause large pricing and hedging errors.

A2 does not hold, but measures exist for incorporating known or stochastic interest rates into the formula, such as that in Merton's Stochastic Interest Rate Model. However, Chance (2003) finds that the Black-Scholes calculations are not highly sensitive to changes in the interest rate, so that incorporating a changing interest rate should not be the main criterion for selecting models, but instead a useful (second order) factor to consider if possible.

The failure of A4, due to commission fees and taxes, is potentially more serious though. Black-Scholes pricing relies on the ability to rebalance the riskless portfolio at every instant. The existence of transactions costs will effectively limit the profitable number of times rebalancing can occur, so leading to errors offset against the transaction costs. This is a potential route of enquiry. But it might be argued that transaction costs will effect all models equally, so that again this may not be the best means of selecting models.

A9 is generally held to be valid owing to the large number of agents and institutions who can eliminate arbitrage opportunities quickly, and due to pricing studies such as that by Klemkosky and Resnick (1979) who tested put-call parity and found that they could not reject it. However, there are problems with the delta hedge employed in the no-arbitrage argument, which are potentially compounded by violations of A1. In reality, the total change in call price is a function of all its determinants, S, τ, r and σ (the strike price K is constant). Implementing a first-order Taylor series expansion,

$$dc = \Delta dS + \frac{1}{2}\Gamma(dS)^2 + \Theta d\tau + \rho dr + \Lambda d\sigma$$

where Γ is the rate of change of Δ with respect to changes in S and Θ, ρ, Λ are sensitivities of c with respect to changes in τ, r and σ

So the Black-Scholes hedge is only a delta neutral hedge, at the very least failing to account for the Γ term. In this case it is only (approximately) riskless instantaneously ($d\tau$ is very small). Also, r and σ must be constant, or changes in r and σ are uncorrelated with changes in c ($\rho = \Lambda = 0$, Θ will always be less than 0 because of unfavourable time decay for European options).

Yet if A1 is violated, the last two terms become significant, unless $\rho = \Lambda = 0$, which is unlikely, because then r and σ should not be determinants of c at all. Bakshi, Cao and Chen (1997, pp. 2034) add that discontinuous share price paths, due to jumps, introduce jump risk which also makes the standard hedge imperfect. In all cases some hedging error exists with the delta hedge, meaning that further pricing and hedging errors will result from implementing Black-Scholes.

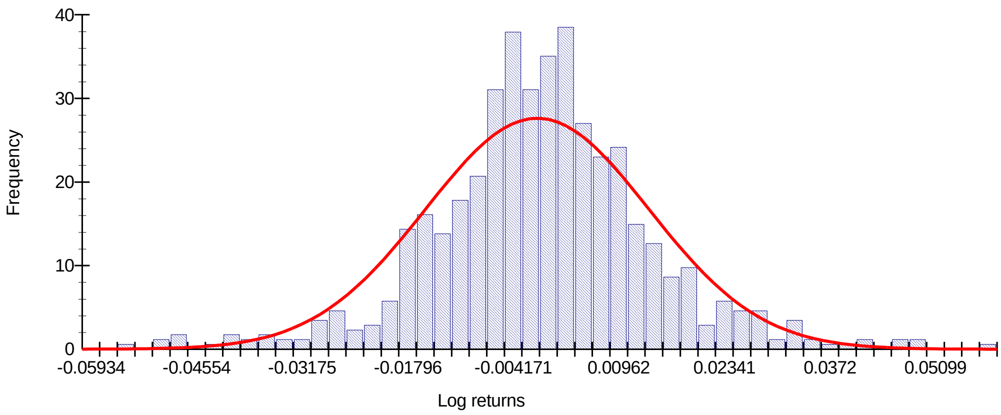
More generally, the failure of A1 has the ability to cause the largest errors. There is significant evidence that GBM does not fit share returns, as will be presented in the next section. Modifying A1 seems to be the main way to improve on Black-Scholes, so models will mainly be chosen to correct for the errors stemming for A1. To conclude then, the thesis will henceforth focus on deviations from GBM, while trying to implement accurate hedges (A9), and correct for transaction costs (A4) and stochastic interest rates (A2) where possible. In order to find better models, the stylised facts which need to be addressed will now be explored.

3B. EMPIRICAL FINDINGS ON THE SHARE PRICE PROCESS

There is a long list of stylised facts which demonstrate that Geometric Brownian Motion is not completely realistic; these are presented in articles by Ghysels, Harvey and Renault (1996, pp. 127-132), Bollerslev, Engle and Nelson (1994), Malone and Rydberg. The focus will be on five of these features, namely 1) leptokurtosis, 2) asymmetry, 3) volatility clustering, 4) leverage effects, and 5) volatility matrices. In line with the hypothetico-deductive methodology of this thesis, a preliminary exploration of share data will be made during the following discussion.

First, the empirical distribution of logged price returns is observed to exhibit fat tails, in that there is a higher probability of a very high or very low return than would be predicted using a normal distribution. At the same time, the curvature of the distribution is greater than under normality, resulting in a higher peak. These features can be seen in Figure 1 (below), which charts the frequencies of daily returns from the FTSE 100 index share over the period 2001-2003 against the theoretical normal probability density function (pdf)[2]. In particular, the kurtosis coefficient of the empirical distribution is 4.7396, which is higher than that of the theoretical normal pdf, which is 3.

Figure 1 - Histogram of log returns and theoretical Normal curve



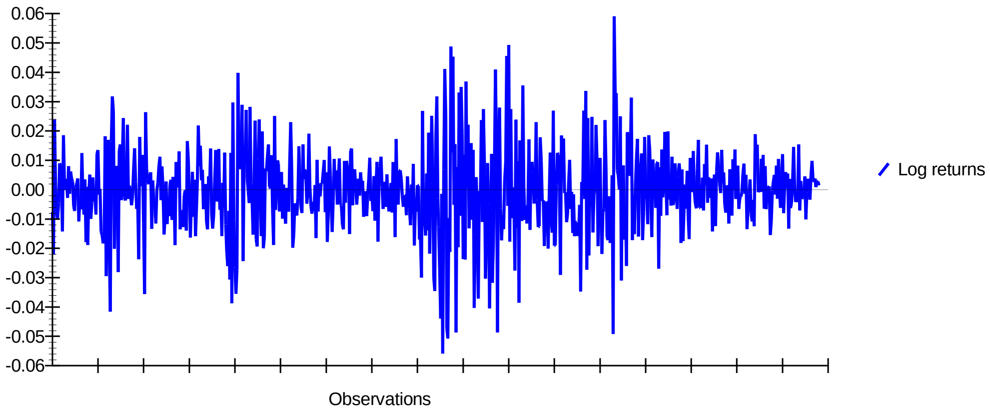
---

[2] This data was obtained from Yahoo UK Finance (http://uk.table.finance.yahoo.com/k?s=^ftse&g=d). It was automatically corrected for dividend issues, so dividends will not be accounted for; the dividend rate *q* is effectively zero so has been omitted from model specifications later.

Second, according to Rydberg, there is evidence that the distribution of share returns is slightly negatively skewed, with the possible explanation that traders react more strongly to negative information; for instance, on Oct 19th, 1987, there was a single large crash in the share market, which required many smaller increases to get back to the previous level. Returning to figure 1, there is a slightly greater mass of observations towards the right side of the distribution, leading to negative skew as exhibited by the skewness coefficient of -0.039199. Carr and Wu (2003, pp. 13) also support negative skewness, though Press (1967, pp. 330-335) and Kon (1984, pp. 161) found positive skewness. In either case, it appears that there is some asymmetry in the return distribution.

Third, large price changes seem to come in bulks – leading to time periods when share return volatility is relatively low and other periods when it is high, as shown in Figure 2 (below), which plots share returns over time for the FTSE 100 index share over the period 2001-2003.

Figure 2 - Log returns over time



Fourth, share price movements have been observed to be negatively correlated with volatility. According to Christie (1982, pp. 408), a rationale consistent with the evidence is that as a firm's income falls, its shareholder value falls, and at the same time, due to fixed costs (e.g. interest payments), its financial leverage rises, increasing its riskiness. However, Christie suggests that leverage alone is too weak to explain empirical asymmetries. Copeland and Weston explain that leverage effects are also disputed by the observations of Blattberg and Gonedes (1974), who found that volatility changes randomly over time, and Rosenberg (1973), who found that it followed an autoregressive scheme (this lends credence to ARCH-type models).

Fifth, volatility is known to vary with time to expiration and moneyness[3], though it should always be constant. Implied volatility, the volatility calculated by inverting the Black-Scholes formula and using observed call prices, is generally found to rise as the option moves further away from being ATM, leading to 'smile' effects, as found by Macbeth and Merville (1979, pp. 279), and Beckers (1980, pp. 667). Furthermore, as time to expiration rises, it is found that implied volatility rises.

In defence of the Black-Scholes model's specification of GBM, the five effects above tend to diminish as the time period under consideration increases due to Aggregational Gaussianity. For example, if weekly rather than daily returns are examined, the weekly return is equivalent to the sum of daily returns[4]. As τ becomes large, the Central Limit Theorem comes into effect, ensuring log returns are Normally distributed and so making GBM more realistic.

Similarly, the Central Limit Theorem also applies for index shares, whose return distribution is a sum of many component shares' distributions, meaning Black-Scholes should be less inaccurate for index options. This convergence can be seen from lower kurtosis and skewness coefficients of the FTSE 100 index (4.7396 and -0.039199) compared with its components', such as Vodafone (7.2954 and 0.15808).

However, it is short time periods which are most important; portfolio managers need to rehedge their positions daily at the very least. Furthermore, index shares are not the only shares in which agents are interested. Therefore models with a more realistic share price process, which can incorporate the stylised facts above, are necessary.

---

[3] 'Moneyness' is the extent to which profits are made from purchasing an option. A formal measure will be defined later. Loosely speaking, when $S < Ke^{-r\tau}$, no profit is made from reselling an option, and the option is out-of-the-money (OTM). When S equals $Ke^{-r\tau}$, no profit or loss is made from resale, and the option is at-the-money (ATM). Profit is made with in-the-money (ITM) options, for which $S > Ke^{-r\tau}$.

[4] Weekly Return = ln(Friday) – ln (Monday) = ln(Friday) – ln(Thursday) + ln(Thursday) – ln(Wednesday) +…-ln(Monday)

Before proceeding to select alternatives to the Black-Scholes model, it is worth noting that the Black-Scholes model is well-established and so enjoys a form of first-mover advantage. According to Hull, traders have become accustomed to the systematic errors produced by the Black-Scholes model; tables of (implied) volatility matrices have been developed in order to be able to price options according to the Black-Scholes formula. Furthermore, the model is relatively simple to implement; there is only one unknown parameter (σ) and the call price is a monotonically increasing function of σ, so that there is a one-to-one mapping between calls and volatilities. Most importantly, its closed-form solution enables it to be computed very quickly.

This implies that any viable replacement for the Black-Scholes model should be simple, fast and accurate, with clear superiority in at least one of these qualities to overcome Black-Scholes' entrenched advantage. In practice though, there is a trade-off between accuracy, which requires more sophisticated modelling, and simplicity and speed. This trade-off will be borne in mind in the next section.

4. ALTERNATIVE MODELS

Alternative mixtures of distributions have been proposed for replacing normality, with some leading to closed-form (Mathematical Finance) alternatives to the Black-Scholes model. Other models which have been proposed generally focus on modifying the volatility process. These alternatives have been able to explain the five stylised facts above with varying degrees of success.

According to Malone, Mandelbrot observed leptokurtosis in commodity returns in 1963, and advanced the general class of stable Paretian (Lévy) distributions as an alternative to the Normal distribution (which is a special case where the tail index parameter $a = 2$). Fama (1965) rejected normality in his tests and found more evidence in support of Mandelbrot's hypothesis. However, most Paretian distributions suffer from having no finite moments of higher order than one, so stable distributions do not exhibit Aggregational Gaussianity. Also, parameters are not easily estimable.

Consequently, a set of non-stable distributions with finite second moments was proposed to model returns, as supported by Press. Press suggested a Poisson mixture of Normal distributions which could yield leptokurtosis and skewness, and by plotting theoretical against empirical cumulative distribution functions, found that such a distribution fitted some of the data well. Praetz (1972, pp. 53-55) suggested the *t*-distribution, which also exhibits leptokurtosis. Using χ^2 tests to test it against Normal, Poisson mixture of Normal and Paretian distributions, Praetz found that it fit the data best. However, Kon found that a discrete mixture of Normal distributions could explain leptokurtosis and skewness better than the t-distribution for common stock and indices. In addition, his model allows for periodic shifts in the mean as well as variance. Madan and Seneta (1990, pp. 521-523) found that a Variance Gamma distribution for share returns outperformed Normal, stable Paretian and the Poisson mixture of Normal distributions under χ^2 tests. Finally, Cox and Ross (1975) generalised GBM to incorporate leverage.

From the various share price processes have come the Finite Moment Logstable Process (LS) from stable Paretian distributions, the Compound Events model from the Poisson mixture of Normal distributions, the Variance Gamma (VG) model from the Variance Gamma process, Merton's Jump Diffusion (MJD) model from combining a jump process and GBM, and the Constant Elasticity of Variance (CEV) model from modified GBM.

Carr and Wu show that their LS model outperforms VG and MJD models, with its main attraction being its ability to fit volatility smiles and term structures simultaneously, rather than requiring calibration to 'flatten' a smile for one maturity and then recalibration for other maturities. It is also relatively parsimonious; LS requires the estimation of two parameters, while VG and MJD models require three and four respectively.

The LS is estimable via the Fast Fourier Transform (FFT); it bypasses the traditional drawback of Paretian distributions by using extreme parameters which give it finite second moments. However, Carr and Wu admit that their model does not capture volatility clustering (since it is a pure jump process), and it suffers in its derivation from not being able to use a riskless asset. The CEV is able to capture leverage, and also outperforms the Black-Scholes model according to studies by Macbeth and Merville, and Beckers. It is also relatively simple, featuring two unknown parameters, though these can be estimated easily using OLS as opposed to the FFT. However, it too cannot explain volatility clustering.

Specific volatility processes have been proposed to overcome the failure of GBM. Engle initially proposed ARCH to model changing uncertainty for inflation, but ARCH-type models applied to option pricing are able to account for most of the stylised facts. In particular, the GARCH model introduced by Bollerslev (1986) is able to account for persistence in the conditional variance, while the EGARCH specification advanced by Nelson (1989) allows ARCH-type models to incorporate skewness (either positive or negative).

With regard to Stochastic Volatility models, Taylor (1986) derived the baseline discrete time version under the Financial Econometrics approach, but continuous-time versions also exist, such as the SV-Random Jump model of Bakshi, Cao and Chen, which has significant explanatory power. Ghysels et al. show that SV models can be created which model all the stylised facts. One such model is Heston's SV (HSV) model. Particularly attractively, this model can also incorporate stochastic interest rates.

Geske's Compound Option model and Rubinstein's Displaced Diffusion model form another approach which makes more use of firm-specific data. However, these models require knowledge of firms' debt and asset structures, which will make evaluating all options more time intensive, but will especially impede fast pricing of options on share indices. Therefore these models will be avoided on the grounds that they are not parsimonious.

In terms of accuracy, the most promising models / starting points for option pricing seem to be LS, the discrete mixture of Normal distributions and SV models. VG, MJD, CEV and EGARCH models also fit some of the stylised facts well. Unfortunately, while Kon's mixture of Normal distributions models share returns well, it has not been developed into an option pricing model. Also, EGARCH requires Monte Carlo implementation which requires significant runtime, a luxury which most traders do not have.

However, the other models can be evaluated quickly. Their underlying processes will now be examined.

5. ADJUSTING THE SHARE PRICE PROCESS

The share price process is key to the development of alternative models. A review of adjustments to the return distribution and the volatility process will demonstrate the origin and characteristics of the models to be explored.

Aggregational Gaussianity applies when the trading interval is fixed and the time horizon becomes large. Due to the large frequency of rebalancing, consideration needs to be given to what happens when the time horizon is fixed but the trading interval becomes small. Then instead of the Central Limit Theorem coming into effect, the Continuous Trading Limit applies.

This means the distribution of share returns should be an infinitely divisible distribution, which is effectively the sum of a normal distribution with independent Poisson variables of different jump sizes (J) and information arrival intensities (λ). The general Markov jump process stemming from such a distribution is:

$$dS = \mu(\Im)dt + dq$$

where $\Im$ is the relevant set of information, which is determined by the particular model specified.

dq is the pure (Poisson) jump process given by:

$$dq = \begin{cases} 0 & \text{with probability } 1 - \lambda(\Im)dt \\ -J & \text{with probability } \lambda(\Im)dt \end{cases}$$

GBM

Further Markov processes are derived as various limits of the process above. The Normal distribution is the most prominent example of an infinitely divisible distribution; then the log returns incorporate a Brownian Motion process (*dB*) which is completely continuous. Under measure *Q*, GBM is:

$$\frac{dS}{S} = rdt + \sigma dB$$

LS

Other infinitely divisible distributions are the gamma, chi-squared and Paretian distributions. More generally, log returns incorporate a Lévy process ($dL_{\alpha,\beta}$), with varying parameters (α, β) which generate the specific infinitely divisible processes. For instance, the LS incorporates maximum negative skewness, so β = -1 and finite moments of all orders (in *S*) are obtained. The LS is:

$$\frac{dS}{S} = rdt + \sigma dL_{a,-1}$$

This degenerates to GBM for α = 2, but otherwise is a pure jump process so is completely discontinuous, and unable to generate volatility clustering.

VG

The VG process is another pure jump process:

$$d\ln S = \{r + \frac{1}{\alpha}\ln[1-(\theta + \frac{\sigma^2}{2})\alpha]\}dt + \theta d\tilde{t} + \sigma dB_{(\tilde{t})}$$

Carr and Madan (2001, pp. 69) show that the process results from evaluating arithmetic Brownian motion, with drift θ and volatility σ, at the random time $d\tilde{t}$, where $d\tilde{t}$ is a subordinated gamma time; it is IID and gamma-distributed $\gamma(dt{:}1, \alpha)$. Consequently, α is the gamma variance rate. Overall, θ and α adjust skewness and kurtosis of returns relative to those for GBM.

MJD

The MJD process blends the continuity of Brownian Motion with the discontinuous nature of pure jump processes;

$$\frac{dS}{S} = (r - \alpha e^{\theta + \frac{h^2}{2}})dt + \sigma dB + (e^k - 1)dJ(\alpha)$$

J(*α*) is a Poisson jump process with intensity *α*. Carr and Wu (2004) explain that jump magnitude *k* is Normally distributed with mean *θ* and variance h^2. The exponential term in the drift component can be shown to be the mean proportional jump given that one jump has already occurred. If *α* = *k* = 0, the process degenerates to GBM.

CEV

As an alternative to altering the stochastic differential or adding jumps, GBM can be modified to incorporate a time varying volatility process, thus:

$$\frac{dS}{S} = rdt + \sigma(S,t)dB$$

The CEV model was one of the first attempts at such a model, imposing $\sigma(S,t) = \delta S^{\beta-2/2}$, leading to:

$$\frac{dS}{S} = rdt + \delta S^{\beta-2/2}dB$$

Here (*β* – 2) is the elasticity of return volatility with respect to the share price. *δ* is the constant of proportionality linking *σ* to *S*, which is identical to a constant share return variance in the special case where *β* = 2; then the diffusion process is GBM and the Black-Scholes formula results from the derivation process sketched in Section 2. A number of further special cases exist. When *β* ≠ 2, heteroskedasticity exists in the share returns, and in particular, when *β* < 2, leverage effects ensue. When *β* = 1, the model is the ‘Linear Price Variance’ or ‘Square Root’ Process.

When $\beta = 0$, the model is the 'Constant Variance' or 'Absolute' Process. Both of these processes possess the advantage that the stock price can reach zero (bankruptcy) with positive probability, another feature not exhibited by GBM.

HSV

A more sophisticated treatment by Heston starts with $\sigma(S,t) = \sqrt{v(t)}$ and dB_1 as the stochastic differential, yielding:

$$\frac{dS}{S} = rdt + \sqrt{v(t)}dB_1$$

$$\text{where } d\sqrt{v(t)} = -\frac{\kappa_v}{2}\sqrt{v(t)}dt + \sqrt{\kappa_v \bar{v}}dB_2$$

Applying Itô's Lemma, a Square Root Process in the volatility is obtained:

$$dv(t) = \kappa_v[\bar{v} - v(t)]dt + \sigma_v\sqrt{v(t)}dB_2$$

The drift term implies that the volatility is mean-reverting (with adjustment speed κ_v), generating persistence in the volatility and therefore clustering effects. When there is negative correlation, ρ, between the two Brownian Motions, leverage effects ensue, and negative skewness results. Furthermore, σ_v represents the volatility of volatility, so that a positive volatility increases the kurtosis of share returns. [Finally, the process becomes GBM when volatility $\sqrt{v(t)}$ becomes a constant.]

6. The Time-Changed Lévy Framework

Proceeding from the share price processes, closed-form solutions are found directly for the Black-Scholes and CEV models, making these the simplest models to evaluate.

A closed-form model also results from the HSV process, which in its derivation, makes use of characteristic functions: the general method of specifying a probability distribution even when closed-form probabilities or moment generating functions do not exist.

In fact, as Carr and Wu (2002) show, the characteristic function is a tool which can be used to describe almost all share price processes, providing the foundation for their 'time-changed Lévy' framework, which can be used to derive many option pricing models, particularly CEV, LS, VG, MJD and HSV (or even BS trivially). This encompassing nature is highly attractive since it provides a means of comparing models more directly, through points of departure in their derivation.

When combined with the Fast Fourier Transform (FFT), the time-changed Lévy framework provides a means of pricing options quickly and perhaps more accurately than BS. It is therefore more attractive than other families of models and techniques, namely discrete-time models which require time-consuming Monte Carlo simulation.

At this point it may be noted that five models are being studied in addition to BS. This constitutes too many models to implement and compare, so the weakest alternative model, CEV, will be dropped; this can explain leverage but is likely to be too simplistic to explain smiles and smirks, in comparison with the richer VG and MJD structures.

As a provisional conclusion to this literature survey, the assumption and relaxation of Geometric Brownian Motion is the key to Black-Scholes and the generation of its alternatives. Four models which stem from different share price processes have been identified as being adequately theoretically consistent with stylised facts on share returns and being relatively parsimonious in terms of the number of parameters requiring to be estimated. Further investigation will proceed to compare these models. The next sections will develop the time-changed Lévy framework, by discussing characteristic functions, and then their applications in FFTs.

7. CHARACTERISTIC FUNCTIONS

The pdf, $f(x)$, or moment generating function, $m_X(u) \equiv \mathrm{E}(e^{uX}) = \int_{-\infty}^{+\infty} e^{ux} f(x)dx$, of a random variable X, completely characterise its behaviour. Unfortunately, these do not always exist in closed form for the log returns of Lévy processes, since their moments can be infinite. An alternative function which embodies all information is the characteristic function, $\varphi_X(u) \equiv \mathrm{E}(e^{iuX}) = \int_{-\infty}^{+\infty} e^{iux} f(x)dx$. This can always be written down since it is bounded for all values of u, as shown in Spanos (1998, pp. 113). Carr and Wu explain that for u real and fixed, the characteristic function is the expected value of the location of a random point on the unit circle.

PURE JUMP PROCESSES

The Lévy-Khintchine Theorem holds for all infinitely divisible processes, and shows that the characteristic function of a Lévy process X_t can be represented as follows:

$$\varphi_{X_t}(u) \equiv \mathrm{E}(e^{iuX_t}) = e^{-t\Psi_x(u)}$$

where $\Psi_x(\lambda)$ is the characteristic exponent or cumulant characteristic function;

$$\Psi_x(u) \equiv -ln[\varphi_x(u)] = -iu\mu + \frac{\sigma^2 u^2}{2} + \int_{\Re-\{0\}} {}_{-\infty}^{+\infty} (1 - e^{iux} + iux\mathbf{1}_{|x|<1})\ell(dx)$$

[**1** is an indicator function.]

Overall, the theorem shows that an infinitely divisible distribution is characterised by a triplet of Lévy characteristics (μ, σ, $\ell(.)$), which represent a linear deterministic drift, a Brownian motion and a pure jump component. This last part is the Lévy measure [$\ell(dx)$], which describes the arrival intensity of jumps of all possible sizes for each component of X.

For finite-activity pure jump processes, such as the MJD model, a finite number of jumps occurs within any finite time interval. Formally, $\int_{\Re-\{0\}} {}_{-\infty}^{+\infty} \ell(dx) = \alpha < \infty$ for such processes. In the case of the MJD model, this integral defines the parameter α.

When the integral is not finite, an infinite-activity pure jump process exists, which permits an infinite number of jumps within any interval. The VG and LS models have such processes. Infinite-activity pure jump processes exhibit different levels of variation; finite variation processes, such as the VG process, travel by a finite aggregate absolute distance in a finite time interval. Infinite variation processes, such as the LS process, do not. A table with characteristic exponents and Lévy measures for finite and infinite-activity processes exists in Carr and Wu (2002, pp. 127).

STOCHASTIC TIME

To model stochastic volatility, Carr and Wu apply a stochastic time change to X_t. They specify a 'random time change' $\{T_t\}$ such that $E(T_t) = t$, i.e. the random variable T_t is an unbiased estimate of t, which is calendar time. Simplifying some of their algebra, $T_t = \int_0^t v(s_-)$, where $v(t)$ is the (stochastic) business activity rate.

Intuitively, T_t represents business time; if business activity rises, there is greater output per day (say), so that business time passes faster than calendar time ($\Delta T_t > \Delta t$), and vice versa. This is conditioned by the requirement that $v(t)$ must be strictly positive to ensure T_t always increases; otherwise business time could move backwards. Overall, with business time speeding up and slowing down randomly, constant volatility in calendar time effectively becomes stochastic in business time.

Finding the characteristic function of log returns then amounts to finding the function for X_t when $t = T_t$, i.e. determining the function for $Y_t = X_{T_t}$. A result in the bond pricing literature establishes that the Laplace transform of T_t evaluated at some ζ, that is $\boldsymbol{L}_{T_t}(\zeta)$, is equivalent to $\mathrm{E}[e^{-\zeta\int_o^t v(s_-)ds}]$. If the characteristic function of Y_t were to evaluate as $\mathrm{E}[e^{-\Psi_x(u)\int_o^t v(s)ds}]$, then it would equal the Laplace transform of T_t evaluated at the characteristic exponent of X.

When T_t and X_t are independent, this follows simply;

$$\varphi_{Y_t}(u) \equiv \mathrm{E}(e^{iuY_{T_t}}) = \mathrm{E}[e^{-T_t\Psi_x(u)}] = \mathrm{E}[e^{\Psi_x(u)\int_o^t v(s)ds}] \equiv \boldsymbol{L}_{T_t}[\Psi_x(u)].$$

Essentially, independence corresponds to time-changing a Lévy process so that its distribution is symmetric about zero; then the resulting characteristic function is real and can be evaluated as above.

Introducing correlation between changes in business time and the Lévy process introduces asymmetry and means the characteristic function has an imaginary part. Then the analysis above only holds if the second and third expectations account for complex values, so they must be taken under a complex-valued measure. This measure, $Q(u)$, is also called the leverage-neutral measure, because correlation between changes in T_t and X_t is caused by leverage in log returns. Overall then,

$$\varphi_{Y_t}(u) \equiv \mathrm{E}^P(e^{iuY_{T_t}}) = \mathrm{E}^{Q(u)}[e^{-T_t\Psi_x(u)}] = \mathrm{E}^{Q(u)}[e^{\Psi_x(u)\int_o^t v(s)ds}] \equiv \boldsymbol{L}_{T_t}^{Q(u)}[\Psi_x(u)].$$

Carr and Wu go on to show that the resulting expression for the characteristic function, which is under the objective measure P, can be transformed to the risk-neutral measure Q by making use of exponential martingales or Esscher transforms (the latter are used by Eberlein, Keller and Prause(1998)).

This results in the following characteristic functions (under Q):

MJD: $\varphi_{Y_t}(u) = \exp\{iu[r - \alpha(e^{\theta + \frac{h^2}{2}} - 1) - \frac{\sigma^2}{2}]t - \frac{\sigma^2 t u^2}{2} + \alpha(e^{iu\theta - \frac{h^2 u^2}{2}} - 1)t\}$

VG: $\varphi_{Y_t}(u) = e^{iurt}[1 - (\theta + \frac{\sigma^2}{2})\alpha]^{-\frac{t}{\alpha}}[1 - (iu\theta - \frac{\sigma^2 u^2}{2})\alpha]^{-\frac{t}{\alpha}}$

LS: $\varphi_{Y_t}(u) = \exp[iu(r + \sigma^2 \sec\frac{\pi\alpha}{2})t - t(iu\sigma)^{\alpha} \sec\frac{\pi\alpha}{2}]$

Incidentally, BS is : $\varphi_{Y_t}(u) = \exp[iu(r - \frac{\sigma^2}{2})t - \frac{\sigma^2 t u^2}{2}]$

AFFINE PROCESSES

Affine-activity rate processes[5], such as Heston's, do not require evaluation at the characteristic exponent. Instead they assume the triplet of Levy characteristics are all affine in the state variables; in the case of the HSV model, these are log returns and volatility. Then the Laplace transform of random time can be shown to be exponential affine in the current activity rate, v_0; $\boldsymbol{L}_{T_t}^{Q(u)}(\zeta) = e^{-b(t)v_0 - c(t)}$.

Specifically, in the HSV model, X_t is taken as a standard Brownian Motion B_t and the business activity rate $v(t)$ as the mean-reverting square root process. Correlation ρ is created via correlation between the Brownian motions driving X_t and $v(t)$. Then carrying out the Laplace transform and solving the ODEs for $b(t)$ and $c(t)$, the following characteristic function is found:

[5] These are also known as Affine Jump Diffusion (AJD) models. Along with HSV, MJD is an AJD.

$$\varphi_{Y_t}(u) = e^{-b(t)v_0 - c(t)}$$

where $b(t) = \dfrac{(1-e^{-\xi.t})u^2}{(\xi + \kappa_v - iu\sigma_v\rho) + (\xi - \kappa_v + iu\sigma_v\rho)e^{-\xi.t}}$

$c(t) =$

$$\frac{\theta_v}{\sigma_v^2}\left\{2\ln\left[\frac{2\xi - (\xi - \kappa_v + iu\sigma_v\rho)(1-e^{-\xi.t})}{2\xi}\right] + (\xi - \kappa_v + iu\sigma_v\rho)t\right\}$$

$\xi^2 = (\kappa_v - iu\sigma_v\rho)^2 + \sigma_v^2 u^2$

$\theta_v = \bar{v}\kappa_v$

8. FAST FOURIER TRANSFORMS

The FFT is a numerical technique which enables fast evaluation of option prices for stochastic processes which do not necessarily yield closed-form solutions, such as stable distributions. As long as the characteristic function for log returns has been found, FFTs can be implemented to price options for a range of strike prices simultaneously.

Most basically, the discrete Fourier transform (DFT) maps a sequence of equally spaced points on a circle, *a*, onto another, *b*. The DFT finds the *k*th element of *b* i.e. b_k, via the formula:

$b_k = \boldsymbol{F}(a_j) = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} e^{i(2\pi/n)jk} a_j$, where $e^{i(2\pi/n)}$ is the *n*th root of unity (multiplying a point by this factor causes it to be rotated by one nth of a full circle).

If *a* were a continuous function, a continuous Fourier transform (CFT) would be required. Now $b_k = \boldsymbol{F}(a_j) = \int_{-\infty}^{+\infty} e^{i(2\pi k)j} a_j dj$.

If a_j were $f(Y_T)$, the pdf of the random variable Y_T, then $\boldsymbol{F}[f(Y_T)] = \int_{-\infty}^{+\infty} e^{i(2\pi k)Y_T} f(Y_T) dY_T$. This is equivalent to the characteristic function of Y_T if $u = 2\pi k$, i.e. $b_k = \varphi_{Y_T}(u)$. To summarise, the CFT of a distribution function is the characteristic function.

This is the same as saying that the inverse CFT of the characteristic function is the distribution function (this is also known as Lévy’s inversion theorem). In this case, $f(Y_T) = \boldsymbol{F}^{-1}[\varphi_{Y_T}(u)] = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iuY_Y} \varphi_{Y_T}(u) du$.

When implementing models in programs, the inverse DFT is used as an approximation for the inverse CFT. This is:

$\boldsymbol{F}^{-1}[\varphi_{Y_T}(u)] = \frac{1}{2\pi}\sum_{u=0}^{n-1} e^{-i(2\pi/n)kY_T}\varphi_{Y_T}(u)$. Now n is chosen to truncate the approximated integral at a suitably large limit.

The (inverse) FFT is a considerably more efficient version of the (inverse) DFT; essentially it exploits repetitions in the DFT to reduce the number of calculations required; with the DFT, runtime is proportional to n^2 (the approximate number of calculations), whereas the (radix-2) FFT has runtime proportional to $n\log_2 n$. A comprehensive guide to the FFT, as well as the DFT, CFT and their interpretation, is given by Cerny (2004, pp. 153-171).

There are a number of different methods which use Fourier transforms to price options. These seem to sort into two main categories:

1. Calculations of probabilities from characteristic functions

Heston derived option prices by using the Lévy/Fourier inversion theorem. Option prices were found using the standard Black-Scholes formula, but with the probability functions no longer being normal but rather a function of the (continuous) inverse transforms of Lévy processes, i.e.:

$$C = S\Pi_1 - Ke^{-r\tau}\Pi_2$$

which has replaced $N(d_1)$ with $\Pi_1 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{\mathrm{e}^{-iu\ln K}\varphi_{Y_T}(u-i)}{iu\varphi_{Y_T}(-i)}\right)du$

and $N(d_2)$ with $\Pi_2 = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left(\frac{\mathrm{e}^{-iu\ln K}\varphi_{Y_T}(u)}{iu}\right)du$

Cerny also sketches an approximation technique for finding probabilities, which has the advantage of using the FFT, thus speeding up calculation. It is an approximation in the sense that, rather than working out the inverse CFT of the characteristic function as above, the density of discretised log returns is found directly by the FFT.

2. Prices through convolution of Fourier transforms

Working with a binomial model for log returns, Cerny shows that the option price can be found through a convolution of the Fourier transform of the option payoff (C_T) and the characteristic function of log returns. In the continuous setting, Lewis (2001) and Carr and Madan (1999) generate their formulae using this same principle, i.e.:

$$C_0 = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \boldsymbol{F}[f(Y_T)]\boldsymbol{F}(C_T)du$$

They also both shift the contour of integration in order to avoid a singularity. However, Lewis leaves the Fourier transform of the payoff in his formula, whereas Carr and Madan simplify their formula to the point where it is only a function of one Fourier transform, thus making it more amenable to being calculated via the FFT.

In more detail, Lewis shows that the Fourier transform of a call payoff is

$-\dfrac{K^{iu+1}}{u^2 - iu}$.

Then $C_0 = S - \dfrac{Ke^{-rT}}{2\pi}\displaystyle\int_{iv-\infty}^{iv+\infty} e^{iuk}\varphi_{Y_T}(-u)\frac{du}{u^2 - iu}$, where $k = \ln\dfrac{S}{K}$

and $0 < v < 1$ (it can be chosen)

By contrast, Carr and Madan show that:

$$C_0(k) = e^{-\alpha k}\boldsymbol{F}^{-1}[\psi_{c_T}(v)] = \frac{e^{-\alpha k}}{\pi}\int_0^{\infty} e^{-ivk}\psi_{c_T}(v)dv$$

Here $k = \ln(K)$, and $\psi_{c_T}(v) = \boldsymbol{F}[c_T(k)]$,

which is shown to be: $\quad = \dfrac{e^{-rT}\varphi_{Y_T}[v-(\alpha+1)i]}{\alpha^2 + \alpha - v^2 + i(2\alpha+1)v}$

α is the damping factor, a value which is chosen to assist in the call price integral. Carr and Madan suggest a value of 1.25.

Overall, since Carr and Madan's approach only requires the evaluation of one integrand (unlike Heston's), and since it is a formal means of implementing the

FFT (unlike the other two approaches), its calculation will be the fastest. A description of the means of implementing the FFT is now provided.

FFT IMPLEMENTATION

In the paper by Carr and Madan, they approximate their call pricing formula with:

$$C_t(k) \approx \frac{e^{-\alpha k}}{\pi} \sum_{j=0}^{N-1} e^{-iv_j k} \psi_{c_T}(v_j)\eta \qquad [N = n \text{, as used before}]$$

Here $v_j = j\eta$, which implies $dv = \eta$. Then the effective upper limit for the integral, a, is equal to $(N-1)\mathrm{dv} = (N-1)\eta$.

The beauty of the FFT is that it returns N values of the call, each for a different k. The distance between each k is λ, i.e. $k_{(l)} = -b + l\lambda$, where $k = 0, \ldots, N-1$.

Since the range of k should be symmetrical about zero, $b = \frac{1}{2}(N-1)\lambda$. Combining all this information,

$$C_t(k) \approx \frac{e^{-\alpha k}}{\pi} \sum_{j=0}^{N-1} e^{-i\lambda\eta jl} e^{-ibv_j} \psi_{c_T}(v_j)\eta$$

This can be seen to be an inverse Fourier transform by taking $\lambda\eta = \frac{2\pi}{N}$. This also shows that the choice of a small η for more accurate integration will cause λ to be larger so that the call prices obtained for different strikes are spread far apart. Paradoxically, this may cause a loss of accuracy in the final call price obtained for a given strike; the call price for a strike which is in between strikes used by the FFT is usually found by (linear) interpolation between two calls with strikes which obey $k_{(l)} = -b + l\lambda$.

Carr and Madan propose the incorporation of weights according to Simpson's Rule to overcome this problem. This means that a more accurate integration can be obtained for a given η, so that λ need not be as large as before. Then the pricing formula becomes:

$$C_t(k) \approx \frac{e^{-\alpha k}}{\pi} \sum_{j=0}^{N-1} e^{-i\frac{2\pi}{N} jl} e^{-ibv_j} \psi_{c_T}(v_j) \frac{\eta}{3}[3 + (-1)^{j+1} - \delta_j]$$

where δ_j is the Kronecker delta function which is 1 when $n = 0$ and 0 otherwise.

In terms of the choice of a and N, which in turn determine η, the Centre for Computational Finance and Economic Agents (CCFEA) in Essex recommend values of 600 and 4096 respectively. Their programs also suggest the replacement of k with $m = \ln(M) = \ln({}^{K}/_{S} e^{rT})$; M is a measure of moneyness, since it takes larger values for OTM options and smaller values for ITM options. This replacement means that the characteristic function of log returns can be used directly without having to correct for the initial level of the stock price (S). However, the call price is now scaled down by a factor of S, so the normalised call price found by this method must be multiplied by S at the end. Then the pricing formula becomes

$$C_t(m) \approx \frac{Se^{-\alpha m}}{\pi} \sum_{j=0}^{N-1} e^{-i\frac{2\pi}{N} jl} e^{-ibv_j} \psi_{c_T}(v_j) \frac{\eta}{3}[3 + (-1)^{j+1} - \delta_j]$$

9. Calibration and Comparison

In line with the hypothetico-deductive methodology outlined at the end of Section 1, results will be generated with a view to examining models' predictive power, data coherence and theoretical consistency. First, the means of estimating parameters will be outlined, and then the means of testing models will be determined.

Means of estimation of parameters

Central to the calculation of prices is the means of estimating parameters, which can occur through historical estimation (maximum likelihood estimation or moment matching) or calibration (selecting model parameters to match observed variables as closely as possible). Most of the discussion in the literature centres on the estimation of volatility, though it can be generalised to all other parameters too.

While all models have different sets of parameters, there is consensus that they should all be estimated by the same method; Macbeth and Merville, and Beckers point out that the same standard of volatility estimate should be used as an input for all compared models, while Bakshi, Cao and Chen claim that estimating all models the same way gives them 'an equal chance' (pp. 2017). The latter also indicate that most practitioners in the option field choose to calibrate their models, and in fact all the three sets of authors mentioned here used implied parameters (those calculated through calibration) when estimating model prices. Bakshi et. al argue that historical estimation requires an excessive amount of data in order to estimate parameters; as Hull points out, a year of data is usually used to estimate the historical volatility of log returns. In contrast, calibration can occur for a single day's data or even less.

However, data requirements for the method of estimation are unimportant if the estimated parameters embody all necessary information. Jiang (2002) explains that while Canina and Figlewski and Lamoureux and Lastrapes found evidence that historical volatility is a better predictor of future volatility, Christensen and

Prabhala found the reverse. Intuitively though, historical estimation seems to be more backward-looking as a method, so may not adequately weight current information (regardless of weighting schemes such as EWMAs). Calibration, by using current prices to back out parameters, is more forward-looking; market prices of options should contain all expectations regarding future price movements if stockmarket efficiency holds. Furthermore, by neglecting the information in current option prices, historical estimation may fail to account for premia embedded therein, so causing inaccuracies.

Overall, it would seem best to calibrate the models to market. This requires programs to generate as good a fit to the actual prices as possible. The fitting method is usually taken to be the minimisation of the sum of squared errors (SSE) arising from the difference between the estimate of some variable and its observed value. If the models being estimated were linear, this would correspond to ordinary least squares, but since they are much more complex, the method is nonlinear least squares. The specific variable chosen (call this the ‘fitting variable’) can take different forms. Different fitting variables suit different circumstances.

For instance, as with Bakshi et al., and Carr and Wu, the squared differences of actual and theoretical option prices are minimised. This enables comparison of models using their pricing errors and implied parameters. In contrast, the CCFEA suggests minimising squared differences between actual implied and theoretical implied volatility (IV). Theoretical IV is found from calculating the option price using a given model, and then calculating the volatility which would produce that price using the Black-Scholes model. Actual IV can be calculated or found from the market. Taking IV as the fitting variable therefore enables an IV surface to be found for each model, which is a more direct method of comparing models than discussing their parameters, some of which will be specific to their model (and therefore incomparable).

This leads onto the criteria for comparing models. No formal set of tests exists, but an attempt will be made to draw up a comprehensive framework.

TESTS OF DATA COHERENCE

In order to ensure the results and conclusions based on them are not invalid, it is necessary to ensure the parameters and prices being calculated are of the right magnitude, and models' behaviours are reasonably as expected.

1. Specifically, minimising pricing errors, the implied parameters can be compared with those in Bakshi et al. and Carr and Madan, which use the same method to generate estimates.

2. Taking IV as the fitting variable, volatility surfaces can be generate to gain insight into models' behaviour over moneyness and maturity. This should agree with previous findings.

TESTS OF PREDICTIVE ABILITY

Friedman (1956) asserts that predictive ability is the strongest test of a model's worth, regardless of assumptions, which need only model the real world sufficiently clearly that the model has predictive power. Models should be able to make predictions beyond the sample data used for their creation, in order to limit data mining.

Tests of predictive ability will be conducted by using the implied parameters for a certain day to generate models' prices on the next day, given the inputs of the next day. Then comparing pricing errors across different moneyness and maturity classes, an idea of the different models' relative strengths can be found.

1. The SSE and average absolute error (AAE[6]) will be the primary means of determining which models are most accurate. Signed error measures (e.g. average pricing error) are less useful in this respect because they may rank a very poor model highly if its large errors conveniently cancel.

[6] This is also known as root mean squared deviation (rmsd)

2. But signed error measures will be used in a secondary capacity, to gain some idea of the direction of bias in estimates. The main statistic will be average pricing error across subsets of the data. This statistic ensures pricing errors are comparable; proportional errors, which will also be reported, may be misleadingly large when option prices are close to zero.

TESTS OF THEORETICAL CONSISTENCY

Testing assumptions is also important; models which have strong theoretical backing are likely to have more widespread usefulness than purely ad hoc specifications. Ideally, tests of all assumptions made for each model should be conducted, though this will detract from the direction of this thesis.

In the context of option pricing, two crucial assumptions (apart from the failure of GBM which has already been accounted for in selecting alternative models) are stockmarket efficiency (A10) and no-arbitrage (A9). These give rise to the Joint Hypothesis problem which potentially affects all the pricing models; it may not be possible to distinguish whether inaccuracy in predictions is due to the model having an incorrect structure or the failure of A10 and A9. In order to try to separate the hypotheses, assumptions can be tested via two sets of tests – tests of market efficiency and tests of no-arbitrage.

1. (Weak-form) market efficiency can be assessed by a test of the random-walk hypothesis for share prices and a Box-Pierce test for serial correlation in returns.
2. A test of put-call parity can establish whether arbitrage opportunities have been fully exploited. The mean of $C - P + Ke^{-r\tau} - S$ will be found for the dataset, and tested for difference with zero using a t-test.

10. DATA AND CALCULATIONS

DATASET

Option data is available from Liffe.com[7]. The dataset procured consists of daily observations on the FTSE 100 index option (European exercise) between 1992 and 2000, which amount to over 2 million lines of data. Among other information, it provides closing option price, strike price, underlying asset value, expiry date, implied volatility, and volume of contracts traded.

Since considerable computational facilities would be required to process so much data, and a shorter dataset is less prone to parameter change due to major crashes such as that in 1997, the dataset will be reduced to the first three months of 2000 (59 524 observations). For the testing of alternative pricing models, only call data will be examined; this reduces the dataset to a more manageable 29 762 observations. In addition, all calls with closing prices or volumes of zero will be excluded as there is effectively no market for these so they would lead to distorted results. This leaves 2864 observations.

Time to expiration will be inputted manually for all options using the expiry date and the convention that non-working days such as weekends and bank holidays are excluded. The number of working days in the year is found to be 246. Trading of shares and options was synchronous, so no provision need be made in this respect.

Data on interest rates is available from the Bank of England at http://www.bankofengland.co.uk/Links/setframe.html. Treasury Bill data was obtained, which will be converted into continuously compounding rates using $r_{\text{continuous}} = ln(1 + r_{\text{annual}})$.

[7] The data was sent directly by an administrator at Liffe.com, so is not generally available on the website. The reduced version has been included with the email version of this thesis.

PROGRAMS

The most suitable programming environment in terms of execution speed, debugging tools and data handling capabilities was provided by Matlab as opposed to Gauss. ***Note to examiner*: *this student has no previous experience with Matlab*. *The language and debugging were self-taught from scratch*.**

Three types of program were used in calculations:

1) Programs which calibrate to minimise in-sample pricing error and generate out-of-sample prices and errors
2) Programs which calibrate to minimise in-sample differences in theoretical and actual IVs and then create a volatility surface
3) Functions shared by all the models, e.g. the iterative IV routine 'impvol'

Programs of type 1) and 2) are included in Appendix A (p. 61). For each model, the pricing program is presented, then the volatility program, and then the FFT implementation, which is shared by the pricing and volatility programs. The only exception is for BS, which does not require a FFT program (it is calculated by the closed-form solution) or a volatility program (this will be explained in the results section). Functions of type 3) are included in Appendix B (p 82).

A very useful repository of information on Matlab implementation of financial techniques exists in the form of the CCFEA website (http://www.essex.ac.uk/ccfea/rc900/rc900.htm). A volatility surface fitting program (the same as HSVV in Appendix B) provided the foundation for programs. It was adapted to the other models in generating their IV surfaces. Its role was smaller in the development of the pricing programs; there its general setup of nonlinear least squares and FFT implementation was used.

11. RESULTS

ASSESSMENT OF DATA COHERENCE

Table 1 shows the mean parameters for the estimated models, with standard deviations in brackets underneath. The first row shows that all the models generate similar log return volatilities which are of a theoretically reasonable size. Examining the parameters for LS, VG and MJD, they all have the same signs as those estimated by Carr and Wu (pp. 29), and in most cases have similar magnitude. Regarding HSV, it can first be noted that negative correlation between Brownian motions exists (-0.77584), which confirms standard theory on leverage. Overall, HSV parameters have the same signs as, and are similar in magnitude to, parameters estimated by Nandi in his estimation of the Heston model, which are used in Shahin (2001, pp. 78).

Table 1 – Average Implied Parameters

n = 2864	Models				
Parameters	BS	LS	VG	MJD	HSV
σ (= $\sqrt{v}$ for HSV)	0.163845 (0.01471)	0.14518 (0.00974)	0.18814 (0.020687)	0.142302 (0.013107)	0.174841 (0.054516)
α	-	1.667355 (0.05048)	0.165343 (0.075435)	0.922363 (0.158975)	-
θ	-	-	-0.38605 (0.167337)	-0.15453 (0.030066)	-
h	-	-	-	0.069327 (0.059319)	-
ρ	-	-	-	-	-0.77584 (0.099898)
κ_v	-	-	-	-	10.96033 (0.679343)
σ_v	-	-	-	-	1.012763 (0.26818)
SSE	1042858	575911	619262.9	544207	452179.5
AAE	19.0821	14.18049	14.70453	13.78464	12.56519

Most importantly, all the parameters are significantly different from zero, which in most cases also demonstrates that the Black-Scholes specification can be improved upon; for example, if all parameters in HSV apart from σ had been found to be zero, HSV would degenerate to BS and there would be no benefit in implementing HSV. In the case of LS, degeneration would occur if α were found to be 2, though it was actually estimated as 1.67.

Moving onto the IV surfaces, the volatility programs could not be run over the whole dataset; the shared program 'impvol' is highly time intensive (it is an iterative search procedure). Instead, the programs were run over the first day. But as can be seen from a comparison of Figure 3 (next page), which is a plot of the whole dataset, and Figure 4, which is a plot of the first day, it may be better to concentrate on one day's data, since Figure 3 is very cluttered. Figure 4 gives a clearer idea of what shape the ideal model should fit, and in the real world, practitioners only try to fit volatility surfaces on a daily basis.

Figure 3 – 3D Plot of calls' Implied Volatility over whole period
[vertical axis is IV, axis on bottom left is maturity, final axis is moneyness]

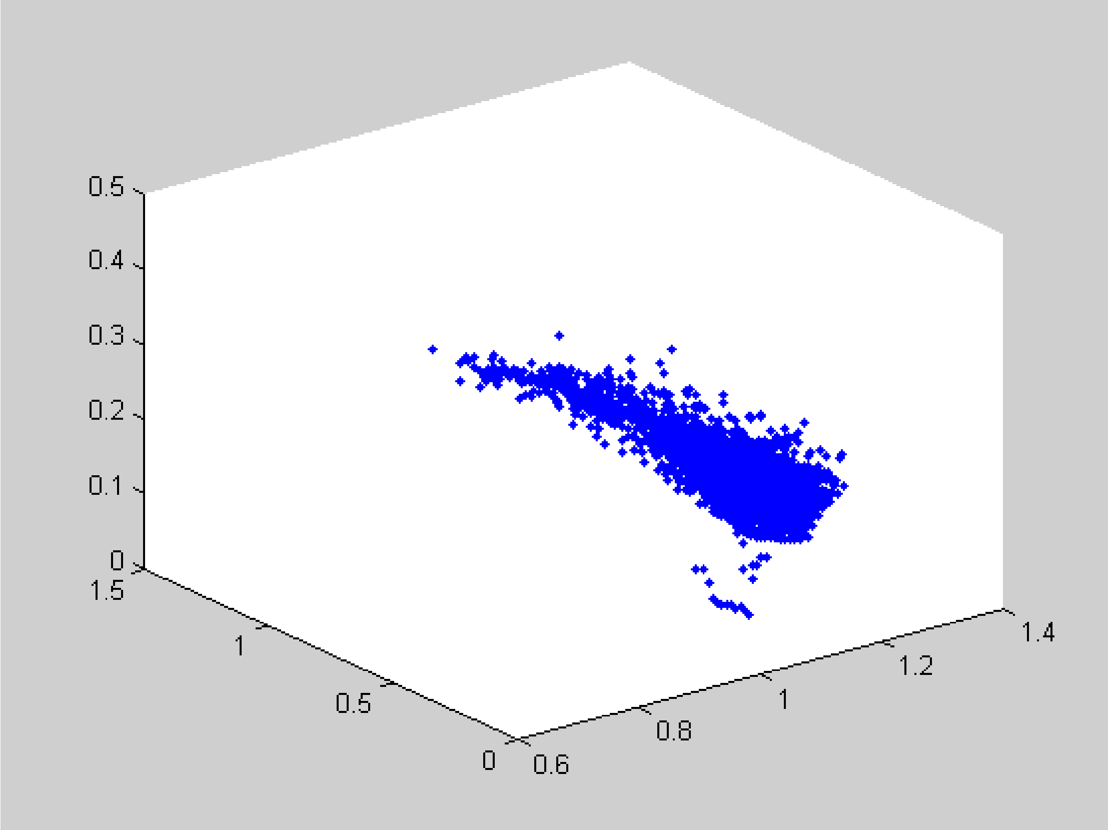
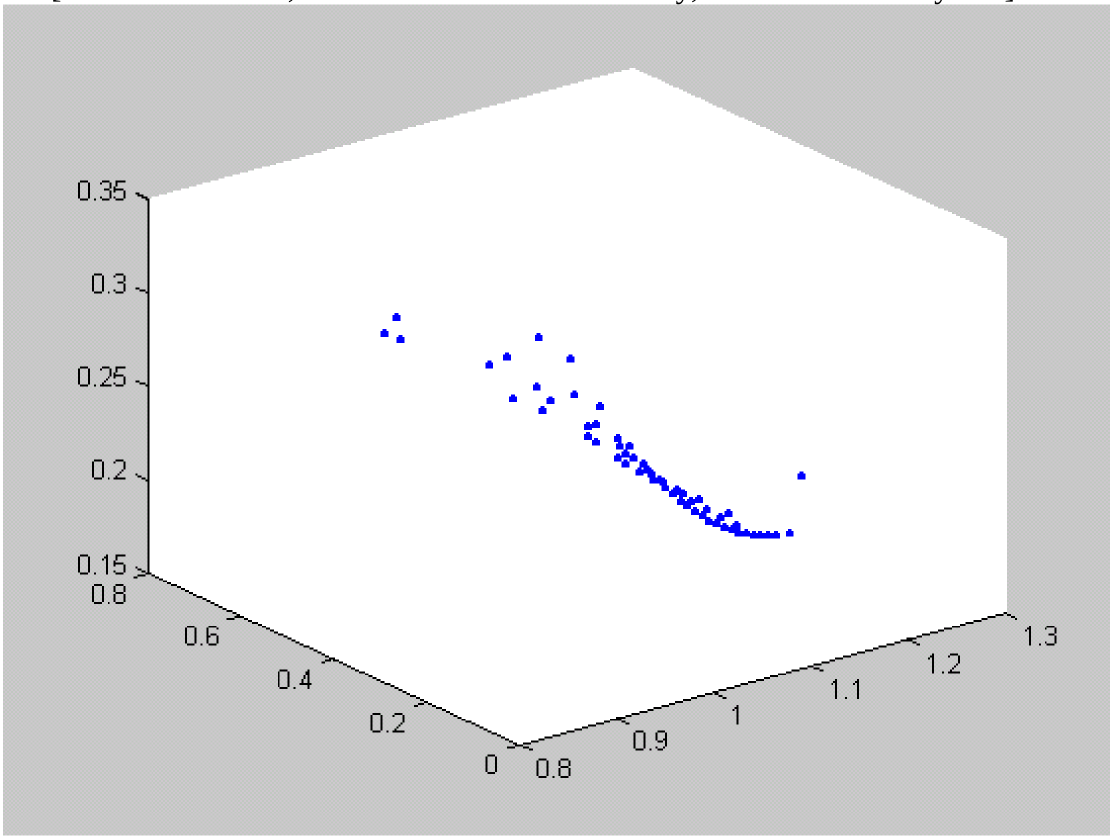


Figure 4 – 3D Plot of calls' Implied Volatility over first day
[vertical axis is IV, axis on bottom left is maturity, final axis is moneyness]



Figures 5-8 (next two pages) show the surfaces for the alternative models, with actual data points represented by the red circles, and implied parameters below the charts[8]. A figure has not been produced for BS, since the constant variance it assumes will result in the trivial form of a flat surface among the data points.

It is worth emphasizing that all the models manage to produce IV surfaces which approximately fit the data, supporting their data coherence, and showing the gain in modelling accuracy which is possible from alternatives to GBM, which would simply produce the aforementioned flat surface.

Among the models it would appear that two observations can be made regarding the complexity (curvature) of the surface and the fit:

1) As will be argued in the predictive ability subsection, LS is able to simultaneously fit across moneyness and maturity, resulting in the vastly more complex shape of its surface than those of VG and MJD. HSV seems to create the second most complex surface, which may be confirmation that "it can induce almost any type of bias to option prices" (Jiang, pp. 64).
2) In terms of fit, HSV seems to be the closest to the actual data, followed by MJD and VG. Despite its complexity, the LS surface actually fits the data worst. This may be a symptom of misspecification.

However, the relative lack of data for long maturities makes such conclusions tentative at best. Analysis over the whole dataset is required to test these conclusions. Comparison via pricing errors rather than informal inspection will make the analysis more rigorous. This will be conducted in the next subsection. Overall though, the models do seem to possess data coherence, though LS is perhaps weakest in this respect.

[8] The main reason for the differences between parameters here and in Table 1 is the different nonlinear least squares problem set up (here the SSE was minimised for IV, while before it was minimised with respect to call price). The parameters can be seen to be of the right signs (apart from *h* for MJD), but the correctness of their magnitudes cannot be readily assessed because no article could be found which minimised SSE for IV.

Figure 5 – Surface plot of (BS) Implied Volatility for LS over first day
[vertical axis is IV, axis on bottom left is maturity, final axis is moneyness]

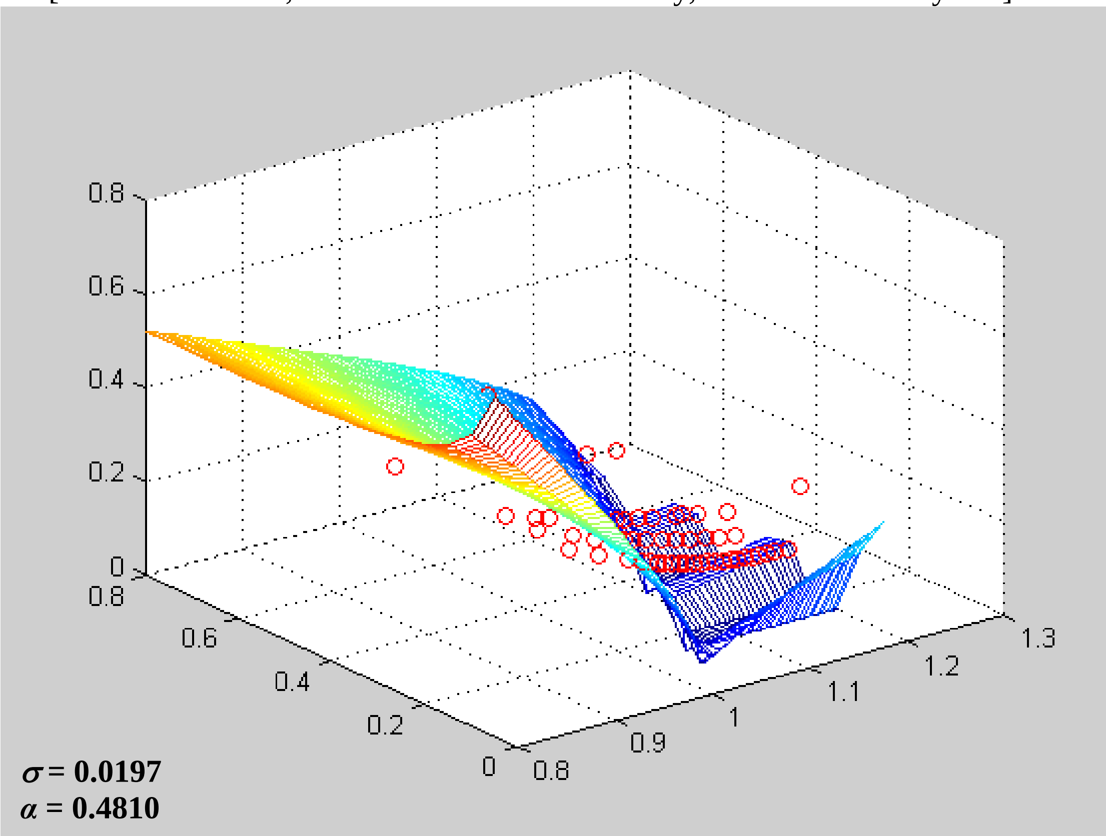


Figure 6 – Surface plot of (BS) Implied Volatility for VG over first day
[vertical axis is IV, axis on bottom left is maturity, final axis is moneyness]

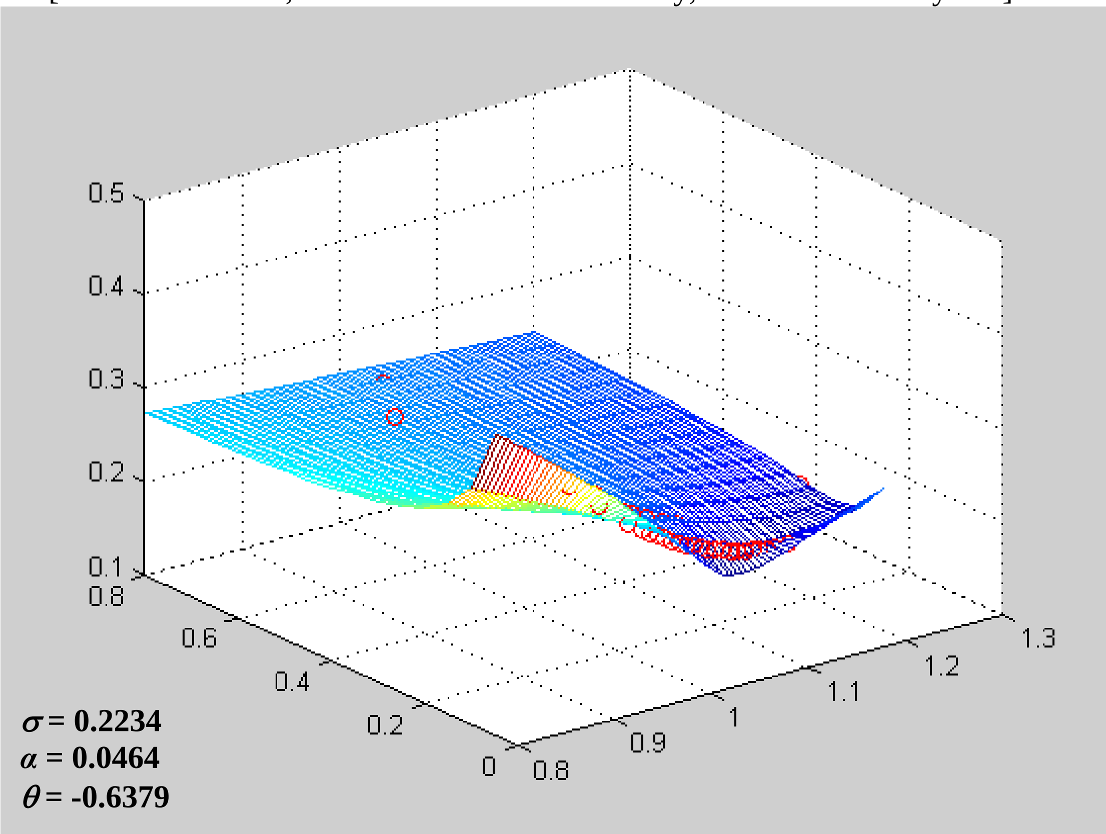


Figure 7 – Surface plot of (BS) Implied Volatility for MJD over first day
[vertical axis is IV, axis on bottom left is maturity, final axis is moneyness]

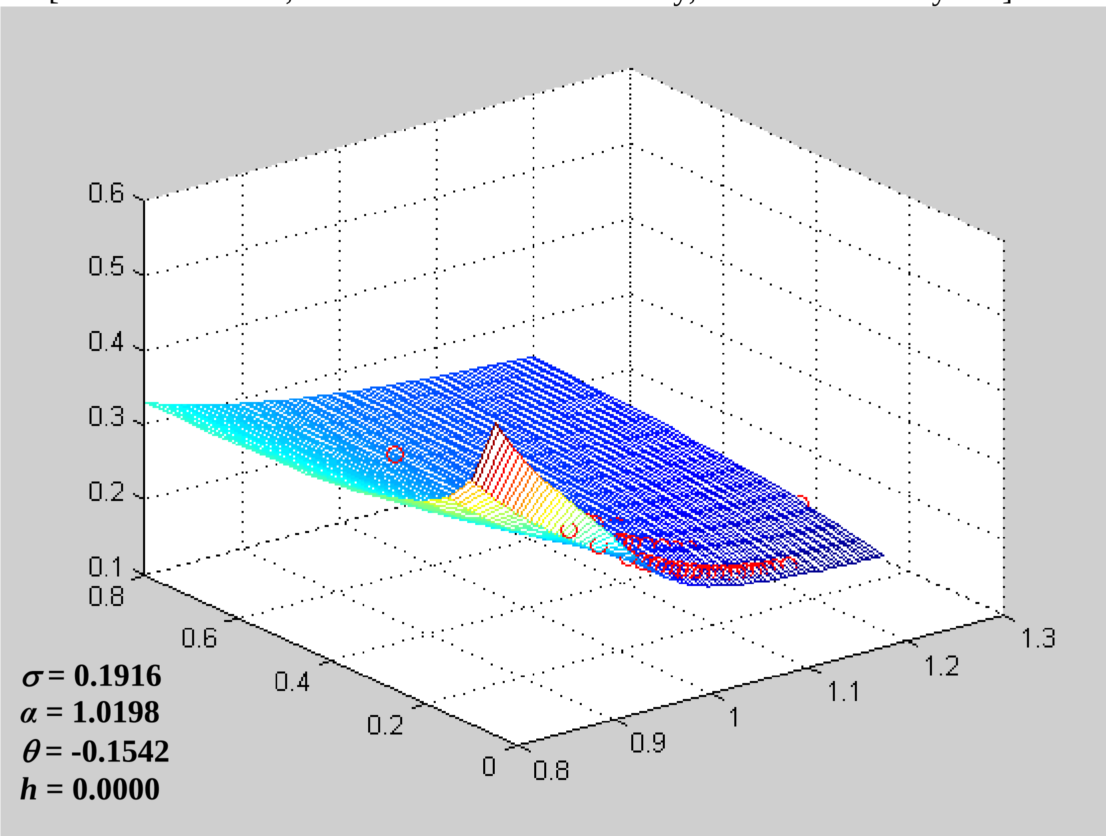
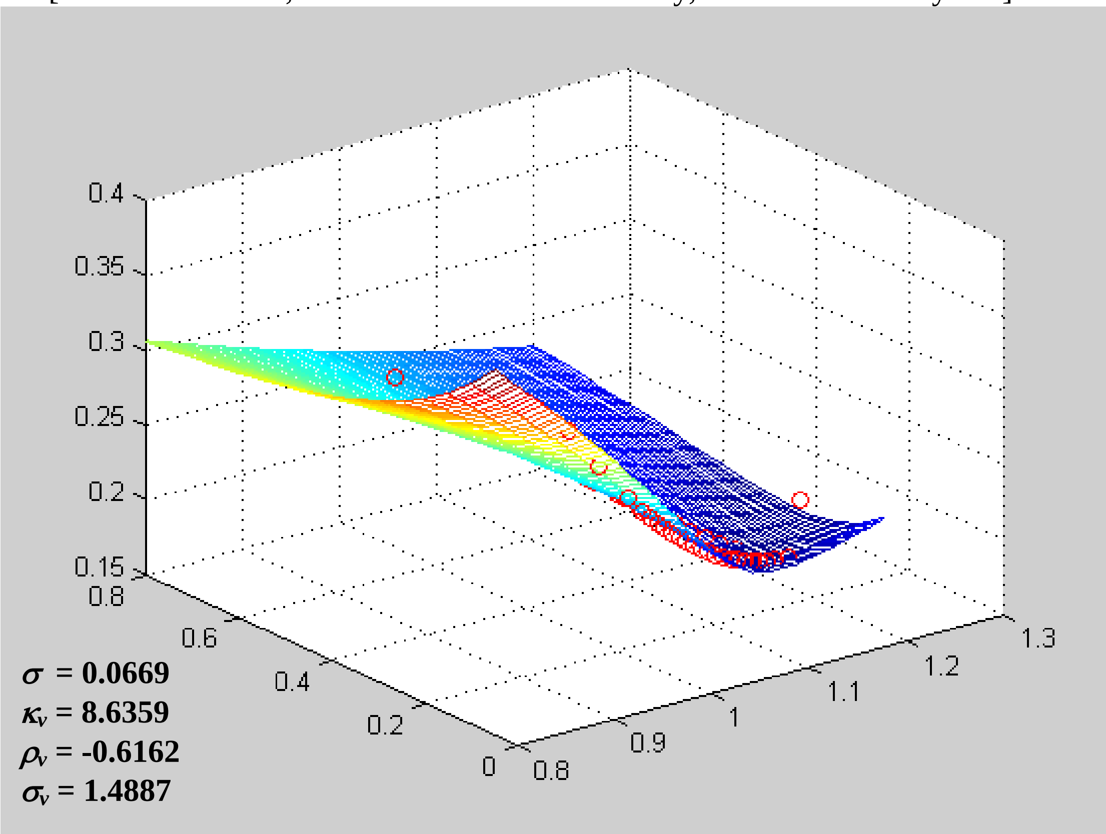


Figure 8 – Surface plot of (BS) Implied Volatility for HSV over first day
[vertical axis is IV, axis on bottom left is maturity, final axis is moneyness]



ASSESSMENT OF PREDICTIVE ABILITY

The statistics for the data as a whole provide the clearest indication of the overall explanatory power of models (these sections have been highlighted), so these will be considered first. The data for specific subsets will then be examined; of these, the nine highlighted subsets are most important, as they contain the bulk of the observations, as shown by Table 2 below, which shows the frequencies for each data class.

Table 2 – Guide to number of observations in each data subset for Tables 3 and 4

	Time to Maturity / days		
Moneyness	<60	[60, 180)	≥180
< 0.8 (Deep ITM)	9	2	0
[0.8, 0.9)	17	14	4
[0.9, 1.0) (ATM)	436	87	37
[1.0, 1.1) (ATM)	1207	282	114
[1.1, 1.2)	320	174	61
> 1.2 (Deep OTM)	5	24	13
Overall	2806		

Concentrating on the unsigned measures of AAE (bottom of Table 3, next page) and SSE (bottom of Table 4, page after next), errors are lowest for the HSV model, which misprices options on average by about £12.65. The next lowest is the MJD, followed by the LS, the VG and finally BS, which misprices options by £19.23. The fact that LS seems to perform better than VG, which has an extra parameter, highlights the possible power a parsimonious model can have relative to more elaborate models. Overall though, HSV seems to minimise absolute error.

Table 3 – Average pricing errors / £

n(observations) = 2806		Time to Maturity / days		
Moneyness	Model	<60	[60, 180)	≥180
< 0.8	BS	42.4	134.8	-
	LS	15.15556	24.95	-
	VG	4.255556	2.65	-
	MJD	12.4	33.15	-
	HSV	1.444444	4.75	-
[0.8, 0.9)	BS	19.20471	44.41071	93.375
	LS	10.03941	3.399286	1.3
	VG	6.525882	-11.6871	-25.15
	MJD	10.73059	2.095	2
	HSV	-3.02235	-9.50643	-3.075
[0.9, 1.0)	BS	2.549718	-10.3823	52.64757
	LS	11.83985	-14.9576	-7.36865
	VG	15.99458	-15.8646	-24.5584
	MJD	9.950289	-14.0972	-7.86757
	HSV	6.151028	-14.8569	-11.2886
[1.0, 1.1)	BS	-2.88855	-15.4712	25.14842
	LS	-1.50316	-7.7829	5.263509
	VG	-2.88269	-4.13898	0.462018
	MJD	-1.90243	-7.14913	6.288509
	HSV	-1.50255	-4.93362	6.573772
[1.1, 1.2)	BS	1.659701	-0.80844	19.35557
	LS	-0.79203	2.95046	22.15033
	VG	1.400759	5.804552	26.13377
	MJD	0.532901	4.063306	24.06393
	HSV	-2.06007	1.407911	22.35098
> 1.2	BS	0.57111	10.48918	29.11577
	LS	-0.96953	10.68866	32.09192
	VG	2.41692	12.69649	37.25923
	MJD	0.147264	11.38912	34.37623
	HSV	-1.24496	2.850684	31.49531
Overall (AAE in brackets)	BS	0.481502 (19.22878)		
	LS	1.022418 (14.28546)		
	VG	1.366611 (14.78616)		
	MJD	0.958454 (13.88278)		
	HSV	-0.08457 (12.64891)		

Table 4 – Relative pricing errors / %

n(observations) = 2806		Time to Maturity / days		
Moneyness	Model	<60	[60, 180)	≥180
< 0.8	BS	2.68792	7.47558	-
	LS	0.94269	1.37301	-
	VG	0.24201	0.1334	-
	MJD	0.7671	1.82672	-
	HSV	0.06734	0.25127	-
[0.8, 0.9)	BS	2.22856	3.83634	6.82233
	LS	1.2411	0.26178	0.06511
	VG	0.87742	-1.065759	-1.889955
	MJD	1.34888	0.17335	0.09475
	HSV	-0.337308	-0.91093	-0.258878
[0.9, 1.0)	BS	5.35876	-2.262918	5.64571
	LS	9.14587	-2.456094	-0.775832
	VG	9.99983	-2.381478	-2.625992
	MJD	7.78969	-2.299082	-0.830999
	HSV	6.6656	-2.334247	-1.195973
[1.0, 1.1)	BS	50.9215	-6.2161	4.37912
	LS	24.3805	-2.708137	1.04304
	VG	44.4667	-1.228086	0.28857
	MJD	38.5589	-2.505778	1.25086
	HSV	11.3268	-1.78297	1.35976
[1.1, 1.2)	BS	80.4291	1.88187	7.55363
	LS	7.20599	6.70192	9.05403
	VG	87.3572	11.4648	10.7967
	MJD	37.1078	10.1369	9.82413
	HSV	-8.975427	-0.121047	8.85443
> 1.2	BS	108.815	44.5562	29.5614
	LS	-44.4302	36.0428	31.6411
	VG	366.506	54.6604	36.655
	MJD	74.0881	40.4335	34.0891
	HSV	-62.26208	8.80623	32.2461
Overall (SSE in brackets)	BS	32.5154 (1037507)		
	LS	13.4150 (572632.7)		
	VG	32.6575 (613477)		
	MJD	23.2359 (540804.5)		
	HSV	4.96448 (448946)		

In terms of the signed measures for the data as a whole, the bottom sections of both Table 3 and Table 4 indicate that HSV performs best by some distance; there is not much to choose between the other models, of which VG now performs worst.

However, the results conflict slightly; the signs of errors disagree in some cases for the same models; e.g. HSV has an (overall) average pricing error of about -£0.08 (underpricing), but has a relative pricing error (overall) of about 4.96% (overpricing). This is due to the cancellation of errors through averaging. The disagreement suggests that while HSV may systematically underprice calls as a whole, it significantly overprices some calls with low values (OTM). This is evident particularly for the subset of the data with maturity less than 60 days and moneyness in the range [1.0, 1.1). Discussion will now focus on each model for the nine subsets in particular.

Black-Scholes: Tables 3 and 4 show that the benchmark model systematically overprices calls. However, consistent with the idea of the volatility smile (across moneyness), BS underprices most / overprices least near the middle of the moneyness range {[1.0, 1.1)}, and overprices more / underprices less on either side of the middle. Table 4 in particular shows that this effect is most pronounced for medium maturity calls, and partially flattens with increasing maturity, which is predicted by aggregational Gaussianity.

These trends largely agree with previous observations, such as those made by Ghysels et al. (pp. 131). However, contrary to convention, the amplitude of the smile is not most for short maturity calls. Furthermore, as Carr and Wu found, volatility smiles need not necessarily flatten with maturity, so that pricing errors may increase in models which allow the onset of aggregational Gaussianity.

It is also surprising that the data indicate pricing errors form a U-shape for given maturities, which contradicts the usual observation of a smirk (upward slope). This effect does not seem to occur in the deep ITM and OTM calls, but for these calls the data is scarcest, so the U-shape may exist for these calls too.

The Alternative Models: The other models exhibit similar behaviour, but for given maturity classes, the smiles are flatter, with minima shifting to calls further ITM. For long maturity calls, the smile has almost completely flattened into an upward slope, as is especially evident in Table 4.

For given moneyness ranges, the U-shaped pattern seen with Black-Scholes also appears, but is less pronounced. For moneyness > 1.1, it has flattened into an upward slope, the expected form.

The result that the error trends are less pronounced for the other models stems from their increased theoretical complexity. Trends in the data which are specific to each model will now be discussed.

Variance Gamma: Of the alternatives to BS, VG performs worst, and in terms of the signed error measures, VG actually performs worse than BS. The only time it seems to outperform the other models is when pricing options with long maturities, for which it produces the widely predicted upward slope most clearly. At the same time though, the errors produced on this slope are among the largest, and VG also produces the largest relative error for a data subset out of all the models; this is 366.5% for the deepest OTM calls of shortest maturity.

Merton's Jump Diffusion: The pricing behaviour of MJD is largely similar to VG, but with considerably smaller errors. This would seem to reflect the improvement in specification possible from including a diffusion component to a pure jump process. Overall, it seems fair to say that, on the basis of this dataset, MJD is superior to VG.

Log Stable: LS pricing errors are almost always smaller than those for VG, and only slightly larger than those for MJD (in fact they are lower in terms of relative percentages). Carr's and Wu's (2004) claim that the model can fit the whole moneyness-maturity surface simultaneously, leading to its superiority over VG and MJD, is borne out partially; average pricing errors are generally smaller for calls which are neither ATM nor medium maturity. Average pricing errors are especially low for long maturity calls, reflecting the fact that the LS process is engineered to avoid the onset of aggregational Gaussianity, which would seem not to hold here.

However, this feature could be criticised as stemming from ad hoc 'reverse engineering' in order to fit the data, rather than resulting from valid economic reasoning; Carr and Wu unrealistically assume infinite variance in the log returns in order to circumvent the Central Limit Theorem.

Heston's Stochastic Volatility: By all measures, HSV misprices calls the least. Only in a few isolated subsets does it fail to have the lowest pricing error, whether average or relative. Also, its errors show the least trendedness, so it fits the actual pricing surface most closely; the average pricing error can be thought of as a Least Squares residual, so since the pricing errors are the least trended / non-random, HSV is least misspecified / it fits the data best.

ASSESSMENT OF THEORETICAL CONSISTENCY

The null hypothesis of non-stationary share prices was not rejected at the 5% significance level, whether or not a time trend was included, thus supporting the idea of a random walk. While not directly confirming stock market efficiency, this result does not rule it out. However, significant positive serial correlation in returns was observed for lags of up to order 10, indicating that share price changes are not random so that the stock market is not even weak-form efficient.

Furthermore, a test of put-call parity revealed that the mean of $C - P + Ke^{-r\tau} - S$, which was -£88.87, was significantly different to zero; the critical region at the (2-tailed) 5% significance level was outside (-1.96, 1.96), while the test statistic was found as -54.0865. So put-call parity does not seem to hold. This failure may be due to the presence of transaction costs, but this is unlikely; estimates of transaction costs given by Klemkosky and Resnick which amount to roughly £40 on average, which is considerably less than the observed mean.

Overall, the joint hypothesis problem persists, reducing the strength of the conclusions in the previous subsection.

12. Conclusion

It was the aim of this thesis to compare the Black-Scholes option pricing model with some of its alternatives, to determine if a better model exists. This analysis was achieved using four prominent alternatives drawn from the encompassing framework of time-changed Lévy processes, which were implemented quickly using the FFT.

Substantial deviations were found between the Black-Scholes model predictions and real-life prices using implied volatility. Less significant deviations arose for the alternative models. This reflects the additional explanatory power the alternative models gained from specifying more realistic share price processes.

Results were always found to be best for HSV, which reflects the importance of Stochastic Volatility as a modelling feature, which in particular is able to generate volatility clustering, unlike the other models tested. It is debatable whether LS or MJD was next best, but due to the poor fit of the LS IV surface (data incoherence) and larger LS pricing errors, MJD would seem to be the next best model. Both models outperform VG, probably due to their additional sophistication. VG is a pure jump, finite variation infinite activity model. Adding a diffusion component leads to an AJD model such as MJD. LS is comparable to MJD even though it is a pure jump process, because it exhibits infinite variation; Carr and Wu argue that this effectively relieves the need for a diffusion component.

So overall, HSV was found to be the best, then MJD, LS, VG and finally BS. This may indicate the superiority of AJD models (HSV and MJD) over pure jump processes (LS and VG), which are in turn superior to GBM.

The parameters and behaviour of models are supported by evidence from other practitioners, apart from the conclusion that MJD outperforms LS, which contradicts the findings of Carr and Wu (2004), but seems valid given the analysis of this thesis. However, tests of two general assumptions of option pricing models general revealed that stockmarket efficiency and no-arbitrage do not hold, so permitting a joint hypothesis problem to persist, and potentially throwing conclusions in doubt.

But these assumption problems may reflect the specific characteristics of the UK; Copeland and Weston show that US studies tend to support stock market efficiency and no-arbitrage. Given that other practitioners have found small pricing errors for option pricing models in the US, assumption problems rather than model misspecification may explain the observed pricing deviations to some extent. Moreover, no other practitioners are known to have performed the tests of theoretical consistency used here for their datasets. So misspecification may exist in many articles.

Nevertheless, the FFT techniques and theoretical understanding required to implement the alternatives to BS were highly challenging and thus constitute a major disadvantage compared to BS, which is much easier to understand. In reality, practitioners may prefer to price options via the ad hoc approach of the empirical determination of the IV surface for each day.

Therefore, an interesting extension would be to examine the IV surface literature (e.g. Cont and Fonseca, 2002) and compare such models to the theoretical models used here. Another extension might also compare the continuous-time models used here with discrete-time models such as EGARCH. Finally, as an additional means of comparison of models (via predictive ability), it would be very useful to evaluate the hedging errors of various portfolio strategies under the models used here; Bakshi et. al provide a derivation of the hedging formulae required under the mean-variance hedging framework (pp. 2034, 2038).

In the meantime, it is the assessment of this thesis that the Black-Scholes model is not the best model. Its disadvantages in terms of accuracy sufficiently outweigh its advantages of parsimony and speed, when compared to the time-changed Lévy models implemented here. Heston’s Stochastic Volatility model has been found to be the clear frontrunner in the time-changed framework, and further research may establish it as the leader in the whole European pricing discipline.

13. REFERENCES

Bakshi, G., Cao, C. and Chen, Z., 1997, Empirical Performance of Alternative Option Pricing Models, *Journal of Finance Vol. LII, No. 5, 2003-2049*

Beckers, S., 1980, The Constant Elasticity of Variance Model and Its Implications For Option Pricing, *Journal of Finance Vol. XXXV No. 3, 661-672*

Black, F. and Scholes, M., 1973, The Pricing of Options and Corporate Liabilities, *Journal of Political Economy 81, 637-659*

Blattberg, R., and Gonedes, N., 1974, A Comparison of the Stable and Student Distributions as Stochastic Models for Stock Prices, *Journal of Business, 244-280*

Bollerslev, T., Engle, R.F. and Nelson D.B., 1994, ARCH Models, in *Engle and Mcfadden – Handbook of Econometrics Vol. IV, Chapter 49*. Elsevier Science B.V.

Bollerslev, T., 1986, Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics 31, 307-327*

Carr, P. and Madan, D.B., 1999, *Option Valuation using the Fast Fourier Transform, Journal of Computational Finance, 2, 61-73*

Carr, P. and Wu, L., 2004, The Finite Moment Logstable Process and Option Pricing, *Journal of Finance, Vol. 58 No. 2, 753-778*

Carr, P. and Wu, L., 2002, Time-changed Lévy processes and option pricing, *Journal of Financial Economics 71, 113-141*

Cerny, A., 2004, *Mathematical Techniques in Finance: Tools for Incomplete Markets*, Princeton University Press

Chance, D.M., 2003, *An Introduction to Derivatives and Risk Management*, South Western College Publishing

Christie, A.A., 1982, The Stochastic Behaviour of Common Stock Variances, *Journal of Financial Economics 10, 407-432*

Cont, R. and Fonseca, J., 2002, Dynamics of Implied Volatility Surfaces, *Quantitative Finance, Vol. 2, 45-60*

Copeland, T.E. and Weston, J.F., 1992, *Financial Theory and Corporate Policy*, Addison-Wesley

Cox, J.C. and Ross, S.A., 1975, The Valuation of Options for Alternative Stochastic Processes, *Journal of Financial Economics 3, 145-166*

Dixit, A.K., and Pindyck, R.S., *1995*, The Options Approach to Capital Investment, *Harvard Business Review May-June, 105-115*

Duan, J-C., 1995, The GARCH Option Pricing Model*, Mathematical Finance Vol. 5 No. 1, 13-32*

Eberlein, E., Keller, U. and Prause K., 1998, *The Journal of Business, Vol. 71, No. 3., 371-405*

Fama, E., 1965, Stock Price Behaviour*, Journal of Business Vol. 38, No. 1, 34-105*

Friedman, M., 1956, *Essays in Positive Economics*, University of Chicago Press

Ghysels, E., Harvey, A.C. and Renault, E., 1996, Stochastic Volatility, in *Maddala and Rao – Handbook of Statistics, Vol.14,* Amsterdam

Harvey, A.C., 1990, *The Econometric Analysis of Time Series,* Pearson Education

Harvey, A.C., 1993, *Time Series Models,* Pearson Education

Jiang, G. J., 2002, Stochastic Volatility and Option Pricing, in *Satchell and Knight – Forecasting Volatility in the Financial Markets*, Butterworth-Heinemann

Klemkosky, R. and Resnick, B., 1979, Put-Call Parity and Market Efficiency, *Journal of Finance, 1141-1155*

Kon, S.J., 1984, Models of Stock Returns – A Comparison*, Journal of Finance, Vol. 39 No. 1, 147-165*

Lewis, A.L., 2001, A Simple Option Formula for General Jump-Diffusion and other Exponential Lévy Processes, *OptionCity.net*: http://www.optioncity.net/pubs/ExpLevy.pdf

Macbeth, J.D. and Merville, L.J., 1980, Tests of the Black-Scholes and Cox Call Option Valuation Models*, Journal of Finance Vol. XXXV, No. 2*

Madan, D.B. and Seneta, E., 1990, The Variance Gamma (V.G.) Model for Share Market Returns*, Journal of Business, Vol. 63, No. 4, 511-524*

Malone, S., 2002, Alternative Price Processes for Black-Scholes: Empirical Evidence and Theory: www.math.duke.edu/vigre/pruv/studentwork/malone.pdf

Mandelbrot, B. B., 1963, New Methods in Statistical Economics*, Journal of Political Economy 71, 421-440*

Merton, R.C., 1998, Applications of Option-Pricing Theory: Twenty-Five Years Later, in *Real Options and Investment Under Uncertainty: Classical Readings and Recent Contributions (by Schwartz, E. S. and Trigeorgis, L.),* MIT Press

Mishkin, F., 2003, *The Economics of Money, Banking and Financial Systems,* Addison Wesley

Nelson, D.B., 1991, Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica, Vol. 59, No. 2., 347-370*

Praetz, P.D., 1972, The Distribution of Share Price Changes, *Journal of Business Vol. 45, No. 1, 637-654*

Press, S.J., 1967, A Compound Events Model for Security Prices, *Journal of Business Vol.40, No. 3, 317-335*

Rosenberg, B., 1973, The Behaviour of Random Variables with Nonstatioary Variance and the Distribution of Security Prices, *manuscript*, *University of California, Berkeley*

Rydberg, T.H., 1999, Realistic Statistical Modelling of Financial Data, to appear in *International Statistical Review.* Also available at: http://www.nuff.ox.ac.uk/users/rydberg/realistic.pdf

Shahin, S. M., 2001,Option Pricing with Stochastic Volatility, *MSc Thesis (Finance), Imperial College of Science, Technology and Medicine*

Spanos, A., 1998, *Probability Theory and Statistical Inference*, Cambridge University Press

Taylor, S.J., 1986, *Modeling Financial Time Series*, Wiley

Thomas, R.L., 2000, *Modern Econometrics,* Addison-Wesley

Tsibiridi, C., 2002, Option Pricing Theory with Stable Distributed Underlyings, *Phd Thesis (Mathematics), Imperial College of Science, Technology and Medicine*

APPENDIX A – PRICING AND VOLATILITY SURFACE PROGRAMS

BS PRICING PROGRAM

```
function xxx=BSC
% program to estimate:
% 1. daily implied volatilities and thereby Black-Scholes prices
% 2. out-of-sample prices based on previous day's implied volatilities

clear;
format compact;

% initialise parameters and output matrices
par = [0.0100];                    % initial parameter estimates
parL= [0.0010];                    % lower bounds
parU= [Inf   ];                    % upper bounds
parF = [];                         % final parameter matrix
global TC ZC ZT ZK ZS ZM;          % intermediate output vectors
AZT = []; AZK = []; AZS = []; AZM = []; ATC = []; AZC = []; BTC = []; % final output vectors

% load data
pts = load('c:/MATLAB6p5/Work/Data.txt');
K = pts(:,1);
S = pts(:,2);
C = pts(:,3);
T = pts(:,4);
r = log(1+pts(:,6)); % Conversion of annualised rates to continuous compounding rates
D = pts(:,7);       % Day index
n = size(D);

% separate data for each day
DI = unique(D);
for jndx=1:size(DI)
   D0 = DI(jndx);

   KD = []; SD = []; CD = []; TD = []; rD = [];    % reset daily inputs
   for indx=1:n
      if (D(indx)==D0)
         KD = [KD; K(indx)];
         SD = [SD; S(indx)];
         CD = [CD; C(indx)];
         TD = [TD; T(indx)];
         rD = r(indx);
      end
   end

   % nonlinear least squares calibration routine
x =…
lsqnonlin(@sumsq,par,parL,parU,optimset('Display','iter','TolFun',0.00001),rD,KD,SD,CD,TD);

   sigma = x; % previous day's parameter

   % concatenate intermediate parameters and outputs for final ones
   parF = [parF; x];
    AZT = [AZT; ZT];
    AZK = [AZK; ZK];
    AZS = [AZS; ZS];
    AZM = [AZM; ZM];
```

```
    ATC = [ATC; TC];
    AZC = [AZC; ZC];

  % generate out-of-sample prices
  if jndx == 1
     BTC = zeros(size(KD),1); % 1st day has no previous implied volatilities to use
  else
     nn = size(KD,1); % nn: rows
     TI = unique(TD);
     FC = [];

     for kndx=1:size(TI)
        T0 = TI(kndx);
        t = T0/246;
        MI = [];
        for lndx=1:nn      % l for llama, not number 'one'
           % take correct maturity times
           if (TD(lndx)==T0)
              MI = [MI; KD(lndx)/SD(lndx)*exp(rD*t)]; % moneyness
              FC = [FC; bs(KD(lndx)/SD(lndx), rD, sigma, t)];    % no adjustment
           end
        end
     end

     BTC = [BTC; FC.*ZS];

  end

end

  % print output
  parF
  [AZT AZK AZS AZM ATC BTC AZC]

%%%%% SUM OF SQUARED DIFFERENCE (THEORETICAL Cs - ACTUAL Cs)^2 %%%%%
function SS = sumsq(par,r,K,S,C,T)
sigma = par;
n = size(K,1); % n: rows
TI = unique(T);
SS = []; % errors

global TC ZC ZT ZK ZS ZM;
ZK = [];
ZT = [];
ZM = [];
FC = [];
ZC = [];
ZS = [];

for jndx=1:size(TI)
  T0 = TI(jndx);
  t  = T0/246;
  for indx=1:n
     % take correct maturity times
     if (T(indx)==T0)
        ZK = [ZK; K(indx)];
        ZM = [ZM; K(indx)/S(indx)*exp(r*t)]; % moneyness
        ZC = [ZC; C(indx)/S(indx)];         % option price
        ZS = [ZS; S(indx)];
        ZT = [ZT; T(indx)];
```

```
        FC = [FC; bs(K(indx)/S(indx), r, sigma, t)];    % no adjustment
      end
   end
end

TC = FC.*ZS; ZC = ZC.*ZS;
SS = TC - ZC;
```

LS PRICING PROGRAM

```
function xxx=LSC
clear;
format compact;

global TC ZT ZK ZS ZM ZC;
% initial values
par = [1.715   0.0100]; % init
parL= [0       0.0010]; % lower bounds
parU= [2       Inf   ]; % upper bounds

% load data
pts = load('c:/MATLAB6p5/Work/Data.txt');
K = pts(:,1);
S = pts(:,2);
C = pts(:,3);
T = pts(:,4);
r = log(1+pts(:,6));
D = pts(:,7);
n = size(D);

AZT = []; AZK = []; AZS = []; AZM = []; ATC = []; AZC = []; BTC = [];
parF = [];      % parameter container

DI = unique(D);
for jndx=1:size(DI)
   D0 = DI(jndx);

KD = []; SD = []; CD = []; TD = []; rD = []; % reset daily secondary parameters

   for indx=1:n
      if (D(indx)==D0)
         KD = [KD; K(indx)];
         SD = [SD; S(indx)];
         CD = [CD; C(indx)];
         TD = [TD; T(indx)];
         rD = r(indx);
      end
   end

% calibration
x=
lsqnonlin(@sumsq,par,parL,parU,optimset('Display','iter','TolFun',0.00001),rD,KD,SD,CD,TD);
alpha = x(1);
sigma = x(2);

    AZT = [AZT; ZT];
    AZK = [AZK; ZK];
    AZS = [AZS; ZS];
    AZM = [AZM; ZM];
    ATC  = [ATC; TC];
```

```
     AZC = [AZC; ZC];
    parF = [parF; x];      % parameter container


if jndx == 1
    BTC = zeros(size(KD),1);
else

nn = size(KD,1); % nn: rows
TI = unique(TD);
SS = []; % errors

for kndx=1:size(TI)
    T0 = TI(kndx);
    t  = T0/246;
     MI = [];
    for lndx=1:nn      % l for llama, not number 'one'
        % take correct maturity times
        if (TD(lndx)==T0)
            MI = [MI; KD(lndx)/SD(lndx)*exp(rD*t)]; % moneyness
        end
    end
    GRD = LSprices(alpha,rD,t,sigma);
    GK = GRD(:,1);
    GC = GRD(:,2);
    GGK = [];
    GGC = [];
    for lndx=1:size(GK)
        if (GK(lndx) > 0.7 & GK(lndx)<1.3)
            GGK = [GGK;GK(lndx)];
           if GC(lndx) < 0
             GC(lndx) = 0;
           end
            GGC = [GGC;GC(lndx)];
        end
    end

     FC = interp1(GGK,GGC,MI);
     SS = [SS; FC];
end

BTC = [BTC; SS.*ZS];

end

end

     [AZT AZK AZS AZM ATC BTC AZC]

%%%%% SUM OF SQUARED DIFFERENCE (THEORETICAL Cs - ACTUAL Cs)^2 %%%%%
function SS = sumsq(par,r,K,S,C,T)
alpha   = par(1);
sigma = par(2);
n = size(K,1); % n: rows
TI = unique(T);
SS = []; % errors

global TC ZT ZK ZS ZM ZC;
ZK = [];
ZT = [];
```

```
ZM = [];
ZC = [];
ZS = [];

for jndx=1:size(TI)
   T0 = TI(jndx);
   t  = T0/246;
    MI = [];
   for indx=1:n
      % take correct maturity times
      if (T(indx)==T0)
         ZK = [ZK; K(indx)];
         MI = [MI; K(indx)/S(indx)*exp(r*t)]; % moneyness
         ZM = [ZM; K(indx)/S(indx)*exp(r*t)];
         ZC = [ZC; C(indx)/S(indx)];          % option price
         ZS = [ZS; S(indx)];
         ZT = [ZT; T(indx)];
      end
   end
   GRD = LSprices(alpha,r,t,sigma);
   GK = GRD(:,1);
   GC = GRD(:,2);
   GGK = [];
   GGC = [];
   for indx=1:size(GK)
      if (GK(indx) > 0.7 & GK(indx)<1.3)
         GGK = [GGK;GK(indx)];
        if GC(indx) < 0
          GC(indx) = 0;
        end
         GGC = [GGC;GC(indx)];
      end
   end

    FC = interp1(GGK,GGC,MI);
    SS = [SS; FC];
end

TC = SS.*ZS; ZC = ZC.*ZS;

SS = TC - ZC;
```

LS VOLATILITY PROGRAM

```
function xxx=LSV
clear;
format compact;

% initial values
par = [1.715   0.0100]; % init
parL= [0       0.0010]; % lower bounds
parU= [2       Inf   ]; % upper bounds
% load data
pts = load('c:/MATLAB6p5/Work/Data1.txt');

K = pts(:,1);
F = pts(:,2);
C = pts(:,3);
T = pts(:,4);
IV = pts(:,5);
```

```
r  = log(1+0.0572); % interest rate
% calibration
par
x= lsqnonlin(@sumsq,par,parL,parU,optimset('Display','iter','TolFun',0.00001),r,K,F,C,T,IV);
x
alpha = x(1);
sigma = x(2);
% presentation of results and plots
% create grid and theoretical IVs
[KK,TT] = meshgrid(0.8:0.005:1.2,0.05:0.05:0.80);
CC = [];
for indx=1:size(TT(:,1))
   t = TT(indx,1); % Take times
   GRD = LSprices(alpha,r,t,sigma);
   GK = GRD(:,1);  % Strikes
   GC = GRD(:,2);  % Option Prices
   FC = interp1(GK,GC,KK(1,:));
   IVs = impvol(KK(1,:),FC,r, t);
   CC = [CC; IVs'];
end
mesh(KK,TT,CC); hold;
plot3(K./F.*exp(r*T/246),T/246,IV,'ro');hold off;

%%%%% SUM OF SQUARED DIFFERENCE (THEORETICAL IVs - ACTUAL IVs)^2
%%%%%
function SS = sumsq(par,r,K,F,C,T,IV)
alpha = par(1);
sigma = par(2);
m = size(K,1); % m: rows
TI = unique(T);
SS = []; % errors
for jndx=1:size(TI)
   T0 = TI(jndx);
   t  = T0/246;
   KI = [];
   CI = [];
   for indx=1:m
      % take correct maturity times
      if (T(indx)==T0)
         KI = [KI; K(indx)/F(indx)*exp(r*t)]; % moneyness
         CI = [CI; C(indx)/F(indx)];          % option price
      end
   end
   GRD = LSprices(alpha,r,t,sigma);
   GK = GRD(:,1);
   GC = GRD(:,2);
   GGK = [];
   GGC = [];
   for indx=1:size(GK)
      if (GK(indx) > 0.7 & GK(indx)<1.3)
         GGK = [GGK;GK(indx)];
         GGC = [GGC;GC(indx)];
      end
   end
   IVs = impvol(GGK',GGC',r, t);
   FV = interp1(GGK,IVs,KI);

   SS = [SS; FV];
end
SS = SS - IV;
```

LS FFT

```
%%%%% FUNCTIONS FOR CHAR. FUNCTION OPTION PRICING
function y = LSprices(alpha,r,t,sigma)
% parameters
a = 600;                % endpoint of char fun grid (0,+a)
%N =  16384;            % number of grid points
N =  4096;              % number of grid points
%N =  200;              % number of grid points
eta = a/N;              % char fun grid size
b = pi/eta;             % +/- bounds for the log-strike grid (-b,+b)
lambda = 2*pi/a;        % log-strike grid size
                        % damping parameter as in Carr-Madan
aa = 1.25;
% create grids
v = (0:N-1) * eta;
m = -b + (0:N-1) * lambda;
% create char fun and invert
h = cfn(v, alpha, r, t, sigma, aa);
h2 = exp(i*b*v) .* h * eta;
g   = fft(h2);
g2  = real( g .* exp(-aa*m) / pi);  % prices based on char fun
y = [exp(m)' g2'];

function y = cfn(th, alpha, r, t, sigma, aa)
% LS characteristic function
% Adjusted to incorporate the Simpson weighting scheme
th1 = th - (aa+1)*i;
omega = r + (sigma^alpha)*sec(pi*alpha/2);

f1 = exp(-r*t) * exp(i*omega*t*th1 - ((i*sigma*th1).^alpha)*t*sec(pi*alpha/2));
f2 = aa^2 + aa - (th.^2) + i*(2*aa+1)*th;
y  = f1 ./ f2;
% Create Simpson's weights
N  = size(th,2);
q1 = (-1).^(1:N);
q2 = eye(1,N);
S  = ( 3 + q1 - q2 )/3;
y  = y .* S;
```

VG PRICING PROGRAM

```
function xxx=VGC
clear;
format compact;

global TC ZT ZK ZS ZM ZC;
% initial values
par = [-0.300   0.1600     0.0100]; % init
parL= [-Inf     0.0010     0.0010]; % lower bounds
parU= [Inf      Inf        Inf   ]; % upper bounds

% load data
pts = load('c:/MATLAB6p5/Work/Data.txt');
K = pts(:,1);
S = pts(:,2);
C = pts(:,3);
T = pts(:,4);
r = log(1+pts(:,6));
```

```
D = pts(:,7);
n = size(D);

AZT = []; AZK = []; AZS = []; AZM = []; ATC = []; AZC = []; BTC =[];
parF = [];     % parameter container

DI = unique(D);
for jndx=1:size(DI)
   D0 = DI(jndx);

KD = []; SD = []; CD = []; TD = []; rD = []; % reset daily secondary parameters

   for indx=1:n
      if (D(indx)==D0)
         KD = [KD; K(indx)];
         SD = [SD; S(indx)];
         CD = [CD; C(indx)];
         TD = [TD; T(indx)];
         rD = r(indx);
      end
   end

% calibration
x= lsqnonlin(@sumsq,par,parL,parU,optimset('Display','iter','TolFun',0.0001),rD,KD,SD,CD,TD);
nu = x(2);
theta = x(1);
sigma = x(3);

    AZT = [AZT; ZT];
    AZK = [AZK; ZK];
    AZS = [AZS; ZS];
    AZM = [AZM; ZM];
    ATC  = [ATC; TC];
    AZC = [AZC; ZC];
   parF = [parF; x];      % parameter container

if jndx == 1
   BTC = zeros(size(KD),1);
else

nn = size(KD,1); % nn: rows
TI = unique(TD);
SS = []; % errors

for kndx=1:size(TI)
   T0 = TI(kndx);
   t  = T0/246;
    MI = [];
   for lndx=1:nn      % l for llama, not number 'one'
      % take correct maturity times
      if (TD(lndx)==T0)
         MI = [MI; KD(lndx)/SD(lndx)*exp(rD*t)]; % moneyness
      end
   end
   GRD = VGprices(theta,rD,t,nu,sigma);
   GK = GRD(:,1);
   GC = GRD(:,2);
   GGK = [];
   GGC = [];
   for lndx=1:size(GK)
```

```
        if (GK(lndx) > 0.7 & GK(lndx)<1.3)
            GGK = [GGK;GK(lndx)];
           if GC(lndx) < 0
              GC(lndx) = 0;
           end
            GGC = [GGC;GC(lndx)];
        end
    end

     FC = interp1(GGK,GGC,MI);
     SS = [SS; FC];
end

BTC = [BTC; SS.*ZS];

end

end

     [AZT AZK AZS AZM ATC BTC AZC]

%%%%% SUM OF SQUARED DIFFERENCE (THEORETICAL Cs - ACTUAL Cs)^2 %%%%%
function SS = sumsq(par,r,K,S,C,T)
theta   = par(1);
nu      = par(2);
sigma   = par(3);
n = size(K,1); % n: rows
TI = unique(T);
SS = []; % errors

global TC ZT ZK ZS ZM ZC;
ZK = [];
ZT = [];
ZM = [];
ZC =[];
ZS =[];

for jndx=1:size(TI)
    T0 = TI(jndx);
    t  = T0/246;
     MI = [];
    for indx=1:n
        % take correct maturity times
        if (T(indx)==T0)
            ZK = [ZK; K(indx)];
            MI = [MI; K(indx)/S(indx)*exp(r*t)]; % moneyness
            ZM = [ZM; K(indx)/S(indx)*exp(r*t)];
            ZC = [ZC; C(indx)/S(indx)];          % option price
            ZS = [ZS; S(indx)];
            ZT = [ZT; T(indx)];
        end
    end
    GRD = VGprices(theta,r,t,alpha,sigma);
    GK = GRD(:,1);
    GC = GRD(:,2);
    GGK = [];
    GGC = [];
    for indx=1:size(GK)
        if (GK(indx) > 0.7 & GK(indx)<1.3)
            GGK = [GGK;GK(indx)];
```

```
        if GC(indx) < 0
          GC(indx) = 0;
        end
         GGC = [GGC;GC(indx)];
      end
   end

    FC = interp1(GGK,GGC,MI);
    SS = [SS; FC];
end

TC = SS.*ZS; ZC = ZC.*ZS;

SS = TC - ZC;
```

VG VOLATILITY PROGRAM

```
function xxx=VGV
clear;
format compact;
% initial values
par = [-0.300   0.1600     0.0100]; % init
parL= [-Inf     0.0010     0.0010]; % lower bounds
parU= [Inf      Inf        Inf   ]; % upper bounds
% load data
pts = load('c:/MATLAB6p5/Work/Data1.txt');
K = pts(:,1);
F = pts(:,2);
C = pts(:,3);
T = pts(:,4);
IV = pts(:,5);
r  = log(1+0.0572); % interest rate
% calibration
par
x= lsqnonlin(@sumsq,par,parL,parU,optimset('Display','iter','TolFun',0.00001),r,K,F,C,T,IV);
x
alpha = x(2);
theta = x(1);
sigma = x(3);
% presentation of results and plots
% create grid and theoretical IVs
[KK,TT] = meshgrid(0.8:0.005:1.2,0.05:0.05:0.80);
CC = [];
for indx=1:size(TT(:,1))
   t = TT(indx,1); % Take times
   GRD = VGprices(theta,r,t,alpha,sigma);
   GK = GRD(:,1);  % Strikes
   GC = GRD(:,2);  % Option Prices
   FC = interp1(GK,GC,KK(1,:));
   IVs = impvol(KK(1,:),FC,r, t);
   CC = [CC; IVs'];
end
mesh(KK,TT,CC); hold;
plot3(K./F.*exp(r*T/246),T/246,IV,'ro');hold off;

%%%%% SUM OF SQUARED DIFFERENCE (THEORETICAL IVs - ACTUAL IVs)^2
%%%%%
function SS = sumsq(par,r,K,F,C,T,IV)
theta   = par(1);
alpha      = par(2);
```

```
sigma   = par(3);
m = size(K,1); % m: rows
TI = unique(T);
 SS = []; % errors
for jndx=1:size(TI)
   T0 = TI(jndx);
   t  = T0/246;
   KI = [];
   CI = [];
   for indx=1:m
      % take correct maturity times
      if (T(indx)==T0)
         KI = [KI; K(indx)/F(indx)*exp(r*t)]; % moneyness
         CI = [CI; C(indx)/F(indx)];          % option price
      end
   end
   GRD = VGprices(theta,r,t,alpha,sigma);
   GK = GRD(:,1);
   GC = GRD(:,2);
   GGK = [];
   GGC = [];
   for indx=1:size(GK)
      if (GK(indx) > 0.7 & GK(indx)<1.3)
         GGK = [GGK;GK(indx)];
         GGC = [GGC;GC(indx)];
      end
   end
   IVs = impvol(GGK',GGC',r, t);
   FV = interp1(GGK,IVs,KI);

   SS = [SS; FV];
end
SS = SS - IV;
```

VG FFT

```
%%%%% FUNCTIONS FOR CHAR. FUNCTION OPTION PRICING %%%%%
function y = VGprices(theta,r,t,alpha,sigma)
% parameters
a = 600;            % endpoint of char fun grid (0,+a)
N =  4096;          % number of grid points
eta = a/N;          % char fun grid size
b = pi/eta;         % +/- bounds for the log-strike grid (-b,+b)
lambda = 2*pi/a;    % log-strike grid size
% damping parameter as in Carr-Madan
aa = 1.25;
% create grids
v = (0:N-1) * eta;
m = -b + (0:N-1) * lambda;
% create char fun and invert
h = cfn(v, theta, r, t, alpha, sigma, aa);
h2 = exp(i*b*v) .* h * eta;
g   = fft(h2);
g2  = real( g .* exp(-aa*m) / pi);  % prices based on char fun
y = [exp(m)' g2'];

function y = cfn(th, theta, r, t, alpha, sigma, aa)
% The characteristic function of the VG model
% Adjusted to incorporate the Simpson weighting scheme
th1 = th - (aa+1)*i;
```

```
f1 = exp(-r*t) * vgcf(th1, theta, alpha, sigma, t, r);
f2 = aa^2 + aa - (th.^2) + i*(2*aa+1)*th;
y  = f1 ./ f2;
% Create Simpson's weights
N  = size(th,2);
q1 = (-1).^(1:N);
q2 = eye(1,N);
S  = ( 3 + q1 - q2 )/3;
y  = y .* S;

function y = vgcf(u, theta, alpha, sigma, t, r)
% Computes the characteristic function for the Variance Gamma model
w  = t*r + t/alpha*log(1 - theta*alpha - .5*alpha*sigma^2);
y0 = exp(i*w*u);
y1 = (1 - i*theta*alpha*u + .5*sigma^2*alpha*(u.^2) ) .^ (-t/alpha);
y  = y0 .* y1;
```

MJD PRICING PROGRAM

```
function xxx=MJDC
clear;
format compact;

global TC ZT ZK ZS ZM ZC;
% initial values
par = [1.0000   -0.010     0.0100     0.1000  ]; % init
parL= [-Inf     -Inf       0.0010     0.0000  ]; % lower bounds
parU= [Inf      Inf        Inf        Inf     ]; % upper bounds

% load data
pts = load('c:/MATLAB6p5/Work/Data.txt');
K = pts(:,1);
S = pts(:,2);
C = pts(:,3);
T = pts(:,4);
r = log(1+pts(:,6));
D = pts(:,7);
n = size(D);

AZT = []; AZK = []; AZS = []; AZM = []; ATC = []; AZC = []; BTC =[];
parF = [];      % parameter container

DI = unique(D);
for jndx=1:size(DI)
   D0 = DI(jndx);

KD = []; SD = []; CD = []; TD = []; rD = []; % reset daily secondary parameters

   for indx=1:n
      if (D(indx)==D0)
         KD = [KD; K(indx)];
         SD = [SD; S(indx)];
         CD = [CD; C(indx)];
         TD = [TD; T(indx)];
         rD = r(indx);
      end
   end

% calibration
x= lsqnonlin(@sumsq,par,parL,parU,optimset('Display','iter','TolFun',0.001),rD,KD,SD,CD,TD);
```

```
alpha   = x(1);
w       = x(2);
h       = x(4);
sigma   = x(3);


     AZT = [AZT; ZT];
     AZK = [AZK; ZK];
     AZS = [AZS; ZS];
     AZM = [AZM; ZM];
     ATC  = [ATC; TC];
     AZC = [AZC; ZC];
    parF = [parF; x];      % parameter container

if jndx == 1
    BTC = zeros(size(KD),1);
else

nn = size(KD,1); % nn: rows
TI = unique(TD);
SS = []; % errors

for kndx=1:size(TI)
    T0 = TI(kndx);
    t  = T0/246;
     MI = [];
    for lndx=1:nn      % l for llama, not number 'one'
        % take correct maturity times
        if (TD(lndx)==T0)
            MI = [MI; KD(lndx)/SD(lndx)*exp(rD*t)]; % moneyness
        end
    end
    GRD = MDprices(alpha,rD,t,w,h,sigma);
    GK = GRD(:,1);
    GC = GRD(:,2);
    GGK = [];
    GGC = [];
    for lndx=1:size(GK)
        if (GK(lndx) > 0.7 & GK(lndx)<1.3)
            GGK = [GGK;GK(lndx)];
           if GC(lndx) < 0
             GC(lndx) = 0;
           end
            GGC = [GGC;GC(lndx)];
        end
    end

     FC = interp1(GGK,GGC,MI);
     SS = [SS; FC];
end

BTC = [BTC; SS.*ZS];

end

end

     [AZT AZK AZS AZM ATC BTC AZC]
```

```
%%%%% SUM OF SQUARED DIFFERENCE (THEORETICAL Cs - ACTUAL Cs)^2 %%%%%
function SS = sumsq(par,r,K,S,C,T)
alpha   = par(1);
w       = par(2);
h       = par(4);
sigma   = par(3);
n = size(K,1); % n: rows
TI = unique(T);
SS = []; % errors

global TC ZT ZK ZS ZM ZC;
ZK = [];
ZT = [];
ZM = [];
ZC =[];
ZS =[];

for jndx=1:size(TI)
    T0 = TI(jndx);
    t  = T0/246;
     MI = [];
    for indx=1:n
        % take correct maturity times
        if (T(indx)==T0)
            ZK = [ZK; K(indx)];
            MI = [MI; K(indx)/S(indx)*exp(r*t)]; % moneyness
            ZM = [ZM; K(indx)/S(indx)*exp(r*t)];
            ZC = [ZC; C(indx)/S(indx)];          % option price
            ZS = [ZS; S(indx)];
            ZT = [ZT; T(indx)];
        end
    end
    GRD = MDprices(alpha,r,t,w,h,sigma);
    GK = GRD(:,1);
    GC = GRD(:,2);
    GGK = [];
    GGC = [];
    for indx=1:size(GK)
        if (GK(indx) > 0.7 & GK(indx)<1.3)
            GGK = [GGK;GK(indx)];
           if GC(indx) < 0
              GC(indx) = 0;
           end
            GGC = [GGC;GC(indx)];
        end
    end

     FC = interp1(GGK,GGC,MI);
     SS = [SS; FC];
end

TC = SS.*ZS; ZC = ZC.*ZS;

SS = TC - ZC;
```

MJD VOLATILITY PROGRAM

```
function xxx=MJDV
clear;
format compact;
% initial values
par = [1.0000   -0.010     0.0100     0.1000  ]; % init
parL= [-Inf     -Inf       0.0010     0.0000  ]; % lower bounds
parU= [Inf      Inf        Inf        Inf     ]; % upper bounds% data load as on 17-Sept-93
pts = load('c:/MATLAB6p5/Work/Data1.txt');
K = pts(:,1);
F = pts(:,2);
C = pts(:,3);
T = pts(:,4);
IV = pts(:,5);
r  = log(1+0.0572); % interest rate
% calibration
par
x= lsqnonlin(@sumsq,par,parL,parU,optimset('Display','iter','TolFun',0.00001),r,K,F,C,T,IV);
x
alpha   = x(1);
w       = x(2);
h       = x(4);
sigma   = x(3);
% presentation of results and plots
% create grid and theoretical IVs
[KK,TT] = meshgrid(0.8:0.005:1.2,0.05:0.05:0.80);
CC = [];
for indx=1:size(TT(:,1))
   t = TT(indx,1); % Take times
   GRD = MDprices(alpha,r,t,w,h,sigma);
   GK = GRD(:,1);  % Strikes
   GC = GRD(:,2);  % Option Prices
   FC = interp1(GK,GC,KK(1,:));
   IVs = impvol(KK(1,:),FC,r, t);
   CC = [CC; IVs'];
end
mesh(KK,TT,CC); hold;
plot3(K./F.*exp(r*T/246),T/246,IV,'ro');hold off;

%%%%% SUM OF SQUARED DIFFERENCE (THEORETICAL IVs - ACTUAL IVs)^2 %%%%%
function SS = sumsq(par,r,K,F,C,T,IV)
alpha   = par(1);
w       = par(2);
h       = par(4);
sigma   = par(3);
m = size(K,1); % m: rows
TI = unique(T);
SS = []; % errors
for jndx=1:size(TI)
   T0 = TI(jndx);
   t  = T0/246;
   KI = [];
   CI = [];
   for indx=1:m
      % take correct maturity times
      if (T(indx)==T0)
         KI = [KI; K(indx)/F(indx)*exp(r*t)]; % moneyness
         CI = [CI; C(indx)/F(indx)];          % option price
      end
```

```
    end
    GRD = MDprices(alpha,r,t,w,h,sigma);
    GK = GRD(:,1);
    GC = GRD(:,2);
    GGK = [];
    GGC = [];
    for indx=1:size(GK)
        if (GK(indx) > 0.7 & GK(indx)<1.3)
            GGK = [GGK;GK(indx)];
            GGC = [GGC;GC(indx)];
        end
    end
    IVs = impvol(GGK',GGC',r, t);
    FV = interp1(GGK,IVs,KI);

    SS = [SS; FV];
end
SS = SS - IV;
```

MJD FFT

```
%%%%% FUNCTIONS FOR CHAR. FUNCTION OPTION PRICING %%%%%
function y = MDprices(alpha,r,t,w,h,sigma)
% parameters
a = 600;            % endpoint of char fun grid (0,+a)
N =  4096;          % number of grid points
%N =  200;          % number of grid points
eta = a/N;          % char fun grid size
b = pi/eta;         % +/- bounds for the log-strike grid (-b,+b)
lambda = 2*pi/a;    % log-strike grid size
% damping parameter as in Carr-Madan
aa = 1.25;
% create grids
u = (0:N-1) * eta;
m = -b + (0:N-1) * lambda;
% create char fun and invert
h = cfn(u, alpha, r, t, w, h, sigma, aa);
h2 = exp(i*b*u) .* h * eta;
g   = fft(h2);
g2  = real( g .* exp(-aa*m) / pi);  % prices based on char fun
y = [exp(m)' g2'];

function y = cfn(th, alpha, r, t, w, h, sigma, aa)
% The characteristic function of the MJD model
% Adjusted to incorporate the Simpson weighting scheme
th1 = th - (aa+1)*i;
f1 = exp(-r*t) * mdcf(th1, alpha, w, h, sigma, t, r);
f2 = aa^2 + aa - (th.^2) + i*(2*aa+1)*th;
y  = f1 ./ f2;
% Create Simpson's weights
N  = size(th,2);
q1 = (-1).^(1:N);
q2 = eye(1,N);
S  = ( 3 + q1 - q2 )/3;
y  = y .* S;

function y = mdcf(u, alpha, w, h, sigma, t, r)
% Computes the characteristic function for the MJD
y  = exp( i*u.*t*(r - alpha*(exp(w + 0.5*h^2) - 1)) - 0.5*(u.^2)*(sigma^2)*t...
     + alpha*t*(exp(i*u.*w - 0.5*(u.^2)*(h^2))-1));
```

HSV PRICING PROGRAM

```
function xxx=HSVC
clear;
format compact;

global TC ZT ZK ZS ZM ZC;
% initial values
par = [0.0100   10.0000    1.0000   -0.6000   0.0100]; % init
parL= [0.0010    0.5000    0.0100   -0.9000   0.0010]; % lower bounds
parU= [Inf      Inf        Inf       0.9000   Inf   ]; % upper bounds
% load data
pts = load('c:/MATLAB6p5/Work/Data.txt');
K = pts(:,1);
S = pts(:,2);

C = pts(:,3);
T = pts(:,4);

r = log(1+pts(:,6));
D = pts(:,7);
n = size(D);

AZT = []; AZK = []; AZS = []; AZM = []; ATC = []; AZC = []; BTC = [];
parF = [];     % parameter container

DI = unique(D);
for jndx=1:size(DI)
   D0 = DI(jndx);

KD = []; SD = []; CD = []; TD = []; rD = []; %DIV = [];   % reset daily secondary parameters

   for indx=1:n
      if (D(indx)==D0)
         KD = [KD; K(indx)];
         SD = [SD; S(indx)];
         CD = [CD; C(indx)];
         TD = [TD; T(indx)];
         rD = r(indx);
      end
   end

% calibration
x= lsqnonlin(@sumsq,par,parL,parU,optimset('Display','iter','TolFun',0.0001),rD,KD,SD,CD,TD);
Vinit  = x(5);
Vbar   = x(1);
kappav = x(2);
thetav = Vbar*kappav;
sigmav = x(3);
rho    = x(4);

    AZT = [AZT; ZT];
    AZK = [AZK; ZK];
    AZS = [AZS; ZS];
    AZM = [AZM; ZM];
    ATC  = [ATC; TC];
    AZC = [AZC; ZC];
   parF = [parF; x];      % parameter container
if jndx == 1
   BTC = zeros(size(KD),1);
```

```
else

nn = size(KD,1); % nn: rows
TI = unique(TD);
SS = []; % errors

for kndx=1:size(TI)
   T0 = TI(kndx);
   t  = T0/246;
    MI = [];
   for lndx=1:nn      % l for llama, not number 'one'
      % take correct maturity times
      if (TD(lndx)==T0)
         MI = [MI; KD(lndx)/SD(lndx)*exp(rD*t)]; % moneyness
      end
   end
   GRD = SVprices(Vinit,rD,t,Vbar,kappav,thetav,sigmav,rho);
   GK = GRD(:,1);
   GC = GRD(:,2);
   GGK = [];
   GGC = [];
   for lndx=1:size(GK)
      if (GK(lndx) > 0.7 & GK(lndx)<1.3)
         GGK = [GGK;GK(lndx)];
        if GC(lndx) < 0
           GC(lndx) = 0;
        end
         GGC = [GGC;GC(lndx)];
      end
   end

    FC = interp1(GGK,GGC,MI);
    SS = [SS; FC];
end

BTC = [BTC; SS.*ZS];

end

 end

    [AZT AZK AZS AZM ATC BTC AZC]

%%%%% SUM OF SQUARED DIFFERENCE (THEORETICAL Cs - ACTUAL Cs)^2 %%%%%
function SS = sumsq(par,r,K,S,C,T)
Vinit  = par(5);
Vbar   = par(1);
kappav = par(2);
thetav = Vbar*kappav;
sigmav = par(3);
rho    = par(4);
n = size(K,1); % n: rows
TI = unique(T);
SS = []; % errors

global TC ZT ZK ZS ZM ZC;
ZK = [];
ZT = [];
ZM = [];
ZC =[];
```

```
ZS =[];

for jndx=1:size(TI)
    T0 = TI(jndx);
    t  = T0/246;
     MI = [];
    for indx=1:n
        % take correct maturity times
        if (T(indx)==T0)
            ZK = [ZK; K(indx)];
            MI = [MI; K(indx)/S(indx)*exp(r*t)]; % moneyness
            ZM = [ZM; K(indx)/S(indx)*exp(r*t)];
            ZC = [ZC; C(indx)/S(indx)];          % option price
            ZS = [ZS; S(indx)];
            ZT = [ZT; T(indx)];
        end
    end
   GRD = SVprices(Vinit,r,t,Vbar,kappav,thetav,sigmav,rho);
    GK = GRD(:,1);
    GC = GRD(:,2);
    GGK = [];
    GGC = [];
    for indx=1:size(GK)
        if (GK(indx) > 0.7 & GK(indx)<1.3)
            GGK = [GGK;GK(indx)];
           if GC(indx) < 0
              GC(indx) = 0;
           end
            GGC = [GGC;GC(indx)];
        end
    end

     FC = interp1(GGK,GGC,MI);
     SS = [SS; FC];
end

TC = SS.*ZS; ZC = ZC.*ZS;

SS = TC - ZC;
```

HSV VOLATILITY PROGRAM

```
function xxx=HSVV
clear;
format compact;
% initial values
par = [0.0100   10.0000    1.0000   -0.6000   0.0100]; % init
parL= [0.0010    0.5000    0.0100   -0.9000   0.0010]; % lower bounds
parU= [Inf      Inf        Inf       0.9000   Inf   ]; % upper bounds
% load data
pts = load('c:/MATLAB6p5/Work/Data1.txt');
K = pts(:,1);
F = pts(:,2);
C = pts(:,3);
T = pts(:,4);
IV = pts(:,5);
r  = log(1+0.0572); % interest rate
% calibration
par
x= lsqnonlin(@sumsq,par,parL,parU,optimset('Display','iter','TolFun',0.00001),r,K,F,C,T,IV);
```

```
x
Vinit  = x(5);
Vbar   = x(1);
kappav = x(2);
thetav = Vbar*kappav;
sigmav = x(3);
rho    = x(4);
% presentation of results and plots
% create grid and theoretical IVs
[KK,TT] = meshgrid(0.8:0.005:1.2,0.05:0.05:0.80);
CC = [];
for indx=1:size(TT(:,1))
   t = TT(indx,1); % Take times
   GRD = SVprices(Vinit,r,t,Vbar,kappav,thetav,sigmav,rho);
   GK = GRD(:,1);  % Strikes
   GC = GRD(:,2);  % Option Prices
   FC = interp1(GK,GC,KK(1,:));
   IVs = impvol(KK(1,:),FC,r, t);
   CC = [CC; IVs'];
end
mesh(KK,TT,CC); hold;
plot3(K./F.*exp(r*T/246),T/246,IV,'ro');hold off;

%%%%% SUM OF SQUARED DIFFERENCE (THEORETICAL IVs - ACTUAL IVs)^2 %%%%%
function SS = sumsq(par,r,K,F,C,T,IV)
Vinit  = par(5);
Vbar   = par(1);
kappav = par(2);
thetav = Vbar*kappav;
sigmav = par(3);
rho    = par(4);
m = size(K,1); % m: rows
TI = unique(T);
SS = []; % errors
for jndx=1:size(TI)
   T0 = TI(jndx);
   t  = T0/246;
   KI = [];
   CI = [];
   for indx=1:m
      % take correct maturity times
      if (T(indx)==T0)
         KI = [KI; K(indx)/F(indx)*exp(r*t)]; % moneyness
         CI = [CI; C(indx)/F(indx)];          % option price
      end
   end
   GRD = SVprices(Vinit,r,t,Vbar,kappav,thetav,sigmav,rho);
   GK = GRD(:,1);
   GC = GRD(:,2);
   GGK = [];
   GGC = [];
   for indx=1:size(GK)
      if (GK(indx) > 0.7 & GK(indx)<1.3)
         GGK = [GGK;GK(indx)];
         GGC = [GGC;GC(indx)];
      end
   end
   IVs = impvol(GGK',GGC',r, t);
   FV = interp1(GGK,IVs,KI);
```

```
    SS = [SS; FV];
end
SS = SS - IV;
```

HSV FFT

```
%%%%% FUNCTIONS FOR CHAR. FUNCTION STOCHASTIC VOL. OPTION PRICING
function y = SVprices(V0,r,t,Vbar,kappav,thetav,sigmav,rho)
% parameters
a = 600;              % endpoint of char fun grid (0,+a)
N =  4096;          % number of grid points
%N =  200;          % number of grid points
eta = a/N;         % char fun grid size
b = pi/eta;        % +/- bounds for the log-strike grid (-b,+b)
lambda = 2*pi/a;    % log-strike grid size
% damping parameter as in Carr-Madan
aa = 1.25;
% create grids
u = (0:N-1) * eta;
x=-b + (0:N-1) * lambda;
% create char fun and invert
h = cfn(u, V0, r, t, kappav, thetav, sigmav, rho, aa);
h2 = exp(i*b*u) .* h * eta;
g   = fft(h2);
g2  = real( g .* exp(-aa*x) / pi);  % prices based on char fun
y = [exp(x)' g2'];

function y = cfn(th, V0, r, t, kappav, thetav, sigmav, rho, aa)
% The characteristic function of the SV model
% Adjusted to incorporate the Simpson weighting scheme
th1 = th - (aa+1)*i;
f1 = exp(-r*t) * cf0(V0, r, t, kappav, thetav, sigmav, rho, th1);
f2 = aa^2 + aa - (th.^2) + i*(2*aa+1)*th;
y  = f1 ./ f2;
% Create Simpson's weights
N  = size(th,2);
q1 = (-1).^(1:N);
q2 = eye(1,N);
S  = ( 3 + q1 - q2 )/3;
y  = y .* S;

function y = cf0(V0, r, t, kappav, thetav, sigmav, rho, u)
% The unadjusted char function
ksi  = sqrt( (kappav - (1 + i*u)*rho*sigmav).^2 - i*u.*(i*u + 1)*sigmav^2 );
% The adjusted char function
ksi0 = sqrt( (kappav - i*u*rho*sigmav).^2 - i*u.*(i*u - 1)*sigmav^2 );
f21  = -2*thetav/sigmav^2* log(1 - ((ksi0 - kappav + i*u*rho*sigmav).*…
(1 - exp(-ksi0*t)))./(2*ksi0));
f22  = -(thetav/sigmav^2)*(ksi0 - kappav + i*u*rho*sigmav)*t;
f24  = (i*u.*(i*u - 1).*(1 - exp(-ksi0*t)))./(2*ksi0 - (ksi0 - kappav + i*u*rho*sigmav).*…
(1 - exp(-ksi0*t)))*V0;
y  = exp(r*t*i*u + f21 + f22 + f24);
```

APPENDIX B – COMMON ROUTINES

```
%%%%% FUNCTIONS FOR IMPLIED VOLATILITIES %%%%%
function y = bs(M, r, sigma, t)
% The simple Black-Scholes European call option Price
d1 = ( -log(M)+( r + .5*sigma^2 )*t ) / sigma/sqrt(t);
d2 = d1 - sigma*sqrt(t);
n1 = normcdf(d1);
n2 = normcdf(d2);
y = n1 - M.*n2*exp(-r*t);

function y = dbs(sigma, M, r, t, c)
y = bs(M, r, sigma, t) - c;

function y = impvol(M, c, r, t)
y = [];
for indx=1:size(M,2)
   MM = M(1,indx);
   cc = c(1,indx);
   yy = fzero(@dbs,0.1,optimset('fzero'),MM,r,t,cc);
   if (yy < 0.01)
      yy = eps;
   end
   y  = [y ; yy];
end
```