Parametric Pricing of Higher Order Moments in S&P500 Options*

G.C.Lim[(a)], G.M.Martin[(b)] and V.L.Martin[(a)]

September 29, 2003

Abstract

A general parametric framework based on the generalized Student t distribution is developed for pricing S&P500 options. Higher order moments in stock returns as well as time-varying volatility are priced. An important computational advantage of the proposed framework over Monte Carlo-based pricing methods is that options can be priced using one-dimensional quadrature integration. The empirical application is based on S&P500 options traded on select days in April, 1995, a total sample of over 100,000 observations. A range of performance criteria are used to evaluate the proposed model, as well as a number of alternative models. The empirical results show that pricing higher order moments and time-varying volatility yields improvements in the pricing of options, as well as correcting the volatility skew associated with the Black-Scholes model.

Key words: Option pricing; volatility smiles and skews; generalised Student t distribution; skewness; kurtosis; time-varying volatility.

JEL classification: C13, G13

(a) Department of Economics, University of Melbourne, Victoria, Australia.

(b) Department of Econometrics and Business Statistics, Monash University, Victoria, Australia.

*This research has been supported by an Australian Research Council Large Grant. The authors would like to thank a Co-Editor and four referees for very constructive and insightful comments on an earlier draft of the paper.

1 Introduction

The Black and Scholes (1973) model represents the most common framework adopted in practice for pricing options. Part of the reason for its popularity is its analytical tractability as the price is simply the mean of a truncated lognormal distribution; see Ingersoll (1987). Two key assumptions underlying the Black-Scholes model are that the distribution of the underlying asset returns is normal and that volatility is constant. However, there is strong empirical evidence that neither assumption is valid; for a review of this literature see Bollerslev, Chou and Kroner (1992). One manifestation of these misspecifications for pricing options is the occurrence of volatility smiles and skews, whereby implied volatility estimates vary across strike prices written on contracts in the same market; see Hull and White (1987), Corrado and Su (1997) and Hafner and Herwartz (2001), amongst others.

A number of alternative frameworks have been proposed in the literature to correct for the misspecification of the Black-Scholes model; see Jackwerth (1999) for a recent review. These frameworks can be classified into three broad categories. The first category involves relaxing the constant volatility assumption. Examples are the deterministic volatility models of Dupire (1994) and Dumas, Fleming and Whaley (1998); the stochastic volatility models of Hull and White (1987), Heston (1993), Bakshi, Cao and Chen (1997), Bates (2000), Chernov and Ghysels (2000) and Pan (2002); and the Generalized Autoregressive Heteroscedasticity (GARCH) models of Engle and Mustafa (1992), Duan (1995, 1999), Heston and Nandi (2000), Hafner and Herwartz (2001) and Bauwens and Lubrano (2002). The second category involves relaxing the normality assumption using either parametric or nonparametric methods. Parametric examples are the lognormal mixture model of Melick and Thomas (1997) and the flexible distributional frameworks adopted in Lim, Lye, Martin and Martin (1998) and Martin, Forbes and Martin (2003). Nonparametric examples include the Edgeworth expansion of Jarrow and Rudd (1982) and Corrado and Su (1997); the Hermite polynomial approximation of Ane (1999); the nonparametric density estimator of Ait-Sahalia (1996), Ait-Sahalia and Lo (1998, 2000) and Ait-Sahalia, Wang and Yared (2001); and the neural network approach of Garcia and Gencay (2000). The third category consists of augmenting the mean returns specification. The most popular form involves the inclusion of a Poisson jump process; see for example, Bakshi, Cao and Chen (1997), Bates (2000) and Pan (2002).

The approach adopted in this paper combines elements of the first two approaches. A parametric risk neutral distribution based on the generalized Student t (GST) distribution of Lye and Martin (1993, 1994) is proposed which accommodates higher order moments in returns distributions. Time varying volatility is modelled by specifying the conditional variance to be a function of the net state return over the life of the option; see Rosenberg and Engle (1997) and Rosenberg (1998). The option is then priced by evaluating the expected value of the discounted payoff of the option contract. The chosen specifications are appealing in that they lead to a computationally efficient procedure for pricing options based on univariate numerical quadrature. This is in contrast to models priced using Monte Carlo methods, which require computing the expectation as an average of a large number of simulation paths. Another advantage of the proposed framework is that a number of existing parametric models are special cases of the GST distribution, including the Black-Scholes model. This means that standard procedures can be adopted to test between competing parametric specifications. Other approaches based on lognormal mixture distributions and Edgeworth expansions are shown to be related to the GST distribution but not directly nested. For these cases other statistical criteria are adopted to test between the competing models. In choosing an appropriate model from the class of models based on Edgeworth expansions, the Jarrow and Rudd (1982) formulation is not adopted as the probability distribution is not guaranteed to be positive over the support. Instead, an alternative option price model based on the semi-nonparametric density of Gallant and Tauchen (1989) is proposed. A final advantage of the proposed framework is empirical, as the GST distribution is shown in general to yield prices of S&P500 options which are superior to other models, in terms of a range of performance measures.

The rest of the paper is structured as follows. Section 2 presents the framework for pricing options using a general parametric family of distributions, based on the GST distribution. Option prices based on this distributional model are referred to as GST option prices. In Section 3 the GST option price model is shown to nest the Black-Scholes model, as well as being related to semi-nonparametric pricing models and a model based on a mixture of lognormals. The shapes of the GST risk neutral probability distributions are investigated, the effects on option pricing examined and the presence of volatility smiles and skews discussed in Section 4. The empirical implications for pricing S&P500 options traded on selected days during April 1995, are presented in Section 5. In evaluating the competing models, five performance measures are adopted based on significance testing, mispricing, forecastability, hedging errors and volatility skew corrections. The key result of the analysis is that the GST model produces option prices that are superior in general to prices produced by all other models considered. A fundamental feature of these empirical results is the importance of modelling skewness in stock returns both to minimize option pricing errors and to establish a consistent framework to price options across the full spectrum of moneyness in a single market. Concluding remarks are contained in Section 6.

2 Parametric Valuation of Options

2.1 General Framework

In this section, a general framework for pricing stock options based on flexible parametric distributions is presented. The distributional model of the stock price generalizes the log-normal distribution which underlies the Black-Scholes option price model, by allowing for higher order moments in returns.

As options on the S&P500 index are European options, the option price model developed here does not allow for early exercise.[1] Consider valuing a European call stock option at time t maturing at time $T = t+n$, where n represents the length of the contract. Defining S_t as the spot price at time t of the stock index, the price of the option with exercise price X, is given as the expected value of its discounted payoff; see Ingersoll (1987) and Hull (2000),

$$F(S_t) = E\left[e^{-r\tau} \max\left(S_T - X, 0\right) |S_t\right], \tag{1}$$

where the conditional expectation $E\left[.|S_t\right]$, is taken with respect to the risk neutral probability measure, $\tau = n/365$ represents the time until maturity expressed as a proportion of a year, and r represents the risk-free interest rate. An alternative way of writing (1) which is more convenient in developing the generalized forms of the risk neutral probability distribution adopted in this paper, is

$$F(S_t) = e^{-r\tau} \int_X^\infty (S_T - X) g(S_T|S_t) dS_T, \tag{2}$$

where $g(S_T|S_t)$ is the risk neutral probability density function of the stock price at the time of maturity, S_T, conditional on the current value, S_t.

[1] However, American style options such as options written on the S&P100 index, could be priced from the framework developed here by using, for example, the upper and lower bounds that characterise the relationships between European and American options; see Melick and Thomas (1997).

In deriving the form of the risk neutral probability distribution, $g(S_T|S_t)$ in (2), the returns of the stock index over the life of the option contract are assumed to be generated as

$$\ln\left(\frac{S_T}{S_t}\right) = \left(r - \frac{\sigma_{T|t}^2}{2}\right)\tau + \sigma_{T|t}\sqrt{\tau}z_T, \tag{3}$$

where $\sigma_{T|t}$ is the annualized conditional volatility process and z_T is a standardized random variable with zero mean and unit variance.[2] In specifying $\sigma_{T|t}$, the formulation of Rosenberg and Engle (1997) and Rosenberg (1998) is adopted,

$$\sigma_{T|t} = \exp\left(\beta_1 + \beta_2 \ln\left(S_T/S_t\right)\right). \tag{4}$$

This specification shows that conditional volatility is stochastic, as it is a function of the future return over the life of the option, $\ln\left(S_T/S_t\right)$. In the case where $\beta_2 > 0$, the relationship between volatility and future return is positive. In the case where $\beta_2 < 0$, there is an inverse relationship between future return and volatility, corresponding to a version of the leverage effect. Setting $\beta_2 = 0$ yields the constant volatility specification which underlies the Black-Scholes model.

The inclusion of future returns in (4) contrasts with volatility specifications based on GARCH, which are backward looking, with volatility being a function of lagged returns. In the context of option pricing, whereby option prices are based on an evaluation of the future evolution of the underlying spot price, equation (4) is a more natural volatility specification. In common with a GARCH-type volatility specification, an additional error term is not introduced, with all randomness deriving from randomness in the asset price itself. This contrasts with a stochastic volatility model, in which the volatility process has its own random innovations, which may or may not be correlated with the innovations to the price process. As is highlighted below, the volatility specification in (4), as well as having a natural interpretation in an option pricing context, also has particular computational advantages compared with the GARCH and stochastic volatility specifications.

In choosing the form of the distribution function of z_T, the adopted distribution needs to be able to capture the well-known empirical feature of nonnormality in stock returns. The distribution adopted here that has these characteristics is the GST distribution introduced

[2] In the case of constant volatility and normal errors, equation (2) can be derived from integrating the stochastic differential equation, $d\ln S_t = \left(r - \frac{\sigma^2}{2}\right)d\tau + \sigma dW$, over the life of the option, where $dW \sim N\left(0, d\tau\right)$. For the more general case of time-varying volatility functions and non-normal errors, the properties of the continuous time process that underlies the model can be identified using the rules of Ait Sahalia (2002).

by Lye and Martin (1993, 1994). This distributional family has been found to be successful in capturing the features of financial returns; see Lim, Lye, Martin and Martin (1998) and Lye, Martin and Teo (1998).[3] Formally the GST distribution is specified as follows. Let w be a GST random variable with mean μ_w, variance σ_w^2 and density given by

$$f(w) = k \exp\left[\theta_1 \tan^{-1}\left(w/\sqrt{\nu}\right) + \theta_2 \ln\left(\nu + w^2\right) + \theta_3 w + \theta_4 w^2 + \theta_5 w^3 + \theta_6 w^4\right], \quad (5)$$

where k is the integrating constant given by

$$k^{-1} = \int \exp\left[\theta_1 \tan^{-1}\left(w/\sqrt{\nu}\right) + \theta_2 \ln\left(\nu + w^2\right) + \theta_3 w + \theta_4 w^2 + \theta_5 w^3 + \theta_6 w^4\right] dw. \quad (6)$$

For the standardized GST variate, $z_T = (w - \mu_w)/\sigma_w$, the density is

$$\begin{aligned} p(z_T) \quad = \quad & k\sigma_w \exp[\theta_1 \tan^{-1}\left(\frac{\mu_w + \sigma_w z_T}{\sqrt{\nu}}\right) + \theta_2 \ln\left(\nu + (\mu_w + \sigma_w z_T)^2\right) \\ & + \theta_3 (\mu_w + \sigma_w z_T) + \theta_4 (\mu_w + \sigma_w z_T)^2 \\ & + \theta_5 (\mu_w + \sigma_w z_T)^3 + \theta_6 (\mu_w + \sigma_w z_T)^4], \end{aligned} \quad (7)$$

where k is the same normalizing constant as defined in (6). Closed form expressions do not exist for k, μ_w and σ_w^2, but these quantities can be computed numerically.

The moments of the GST distribution exist as long as the parameter on the highest even-order term is negative. Hence, with reference to (7), imposing the restriction $\theta_6 < 0$ ensures the existence of all moments of the distribution. The term $\exp\left[\theta_4 (\mu_w + \sigma_w z_T)^2\right]$ corresponds to the kernel of a normal density, with $\theta_4 = -0.5$. The power term, $\left(\nu + (\mu_w + \sigma_w z_T)^2\right)^{\theta_2}$, is a generalization of the kernel of a Student t density and controls the degree of kurtosis in the distribution, along with θ_4 and θ_6. The parameters θ_1, θ_3 and θ_5 control the odd moments of the distribution, including skewness. The role of the parameters of (7) in controlling the shape of the returns distribution is illustrated in Section 3.

The risk neutral probability density function $g(S_T|S_t)$, is derived from the returns distribution $p(z_T)$ in (7), via

$$g(S_T|S_t) = |J|\, p(z_T), \quad (8)$$

[3] Other flexible parametric frameworks can also be used. For example, Martin, Forbes and Martin (2003) use a combination of the GST model and the distributional model of Fernandez and Steele (1998).

where J is the Jacobian of the transformation from z_T to S_T, given by

$$\begin{aligned} J &= \frac{dz_T}{dS_T} \\ &= \frac{1}{S_T \sigma_{T|t} \sqrt{\tau}} \left[1 + \beta_2 \sigma^2_{T|t} \tau - \beta_2 \left(\ln \left(S_T / S_t \right) - \left(r - \frac{\sigma^2_{T|t}}{2} \right) \tau \right) \right], \end{aligned} \tag{9}$$

and $\sigma^2_{T|t}$ is defined in (4).

Stock options can be priced by using (7) to (9) in (2). This formulation expands the Black-Scholes pricing framework as now both kurtosis and skewness in stock returns, as well as conditional volatility, are all priced in the stock option. Apart from some special cases, the integral in (2) needs to be computed numerically. In contrast with other option evaluation methods however, this computation occurs via a straight-forward application of one-dimensional numerical quadrature. In particular, the specification of volatility as a function of S_T enables a closed form solution for $g(S_T|S_t)$ to be derived, meaning that the augmentation of a nonnormal distributional assumption with time-varying volatility involves no additional computational complexity.[4]

Option prices based on (7) to (9) are referred to hereafter as GST prices. In the empirical analysis, three variants of the GST model are investigated:

$$\begin{array}{ll} \text{GST-1:} & \theta_1 \neq 0;\ \theta_2 = -\left(\frac{1+\nu}{2}\right);\ \theta_3 = 0;\ \theta_4 = -0.5;\ \theta_5 = \theta_6 = 0 \\ \text{GST-2:} & \theta_1 = 0;\ \theta_2 = -\left(\frac{1+\nu}{2}\right);\ \theta_3 \neq 0;\ \theta_4 = -0.5;\ \theta_5 = \theta_6 = 0 \\ \text{GST-3:} & \theta_1 = 0;\ \theta_2 = -\left(\frac{1+\nu}{2}\right);\ \theta_3 \neq 0;\ \theta_4 \neq 0;\ \theta_5 \neq 0;\ \theta_6 = -0.25. \end{array} \tag{10}$$

The choice of value for θ_4 in GST-1 and GST-2 is motivated by the usual normalization adopted in the normal distribution. The choice of value for θ_6 in GST-3 also serves as a normalization. Before considering the properties of the GST option pricing model in more detail, its relationship with other option pricing models is discussed in the following section. Three alternative models are discussed. The first is the Black-Scholes model, which is based on the assumption of normal returns and constant volatility. The second is based on semi-nonparametric distributions, constructed from an augmentation of a normal distribution to

[4] The stochastic volatility model of Heston (1993) produces a closed form solution for the option price only under the assumption of conditional normality. Even then, the solution is analytical only up to two one-dimensional integrals in the complex plane. Further, estimation of any stochastic volatility model via observed option prices is extremely computationally intensive due to the presence of the latent volatilites; see, for example, Chernov and Ghysels (2000) and Forbes, Martin and Wright (2003). The same point applies to the stochastic volatility model with jumps investigated by Bakshi, Cao and Chen (1997), Bates (2000) and Pan (2002). The GARCH option price of Heston and Nandi (2000) is similar in nature to the specification in Heston (1993), again based on conditional normality. The augmentation of GARCH with a nonnormal conditional distribution in Hafner and Herwartz (2001), Bauwens and Lubrano (2002) and Martin, Forbes and Martin (2003) entails the use of Monte Carlo simulation for the option price evaluation.

allow for higher order moments. The third approach also captures higher order moments, but via a mixture of lognormals.[5]

3 Relationships with Other Models

3.1 Black-Scholes Option Pricing

The Black-Scholes option price model is based on the assumption that returns are normally distributed. From (7), normality is achieved by imposing the restrictions

$$\theta_1 = \theta_2 = \theta_3 = 0;\ \theta_4 = -0.5;\ \theta_5 = \theta_6 = 0, \tag{11}$$

thereby yielding the standard normal probability density function

$$p(z_T) = ke^{-0.5z_T^2}, \tag{12}$$

with

$$k = \frac{1}{\sqrt{2\pi}},$$

as now $\mu_w = 0$ and $\sigma_w^2 = 1$. Using (12) in (8) gives the risk neutral probability density as

$$g(S_T|S_t) = |J| \exp\left[-\frac{1}{2}\left(\frac{\ln\left(S_T/S_t\right) - \left(r - \frac{\sigma_{T|t}^2}{2}\right)\tau}{\sigma_{T|t}\sqrt{\tau}}\right)^2\right], \tag{13}$$

where J is given by (9) and $\sigma_{T|t}$ by (4).

The other assumption underlying the Black-Scholes model is that volatility is constant over the life of the contract. By setting $\beta_2 = 0$ in (4), (13) simplifies to the lognormal density

$$g(S_T|S_t) = \frac{1}{S_T \exp\left(\beta_1\right)\sqrt{\tau}} \exp\left[-\frac{1}{2}\left(\frac{\ln\left(S_T/S_t\right) - \left(r - \frac{\exp(\beta_1)}{2}\right)\tau}{\exp\left(\beta_1\right)\sqrt{\tau}}\right)^2\right]. \tag{14}$$

Using (14) in (2), the price of the option is

$$F(S_t) = e^{-r\tau}\int_X^\infty \frac{(S_T - X)}{S_T \exp(\beta_1)\sqrt{2\pi\tau}} \exp\left[-\frac{1}{2}\left(\frac{\ln\left(S_T/S_t\right) - \left(r - \frac{\exp(\beta_1)}{2}\right)\tau}{\exp(\beta_1)\sqrt{\tau}}\right)^2\right] dS_T, \tag{15}$$

[5] Comparisons with other classes of option pricing models, as, for example, the stochastic volatility/random jump models mentioned earlier, are beyond the scope of this paper. However such comparisons would be an interesting area of future research.

which is equivalent to the discounted value of the mean of a truncated lognormal distribution; see Ingersoll (1987). For this case an analytical solution exists and is given by the standard Black-Scholes stock option pricing equation,

$$F(S_t) = BS = S_t N(d_1) - Xe^{-r\tau}N(d_2), \tag{16}$$

where

$$\begin{aligned} d_1 &= \frac{\ln(S_t/X) + \left(r + \frac{\exp(2\beta_1)}{2}\right)\tau}{\exp(\beta_1)\sqrt{\tau}} \\ d_2 &= \frac{\ln(S_t/X) + \left(r - \frac{\exp(2\beta_1)}{2}\right)\tau}{\exp(\beta_1)\sqrt{\tau}}. \end{aligned}$$

3.2 Semi-Nonparametric Models

The class of semi-nonparametric option pricing models discussed here are based on an augmentation of the normal returns density through the inclusion of higher order terms. Jarrow and Rudd (1982) were the first to adopt this approach, which has more recently been implemented by Corrado and Su (1997) and Capelle-Blancard, Jurczenko and Maillet (2001).

To show the relationship between the GST and the Jarrow-Rudd option pricing models, consider expanding the GST density in an Edgeworth expansion around the normal density. Letting $p(z_T)$ represent the GST density with distribution function $P(z_T)$, and $n(z_T)$ represent the normal density with distribution function $N(z_T)$, the Edgeworth expansion is

$$\begin{aligned} p(z_T) &= n(z_T) - \frac{(\kappa_1(P) - \kappa_1(N))}{1!}\frac{dn(z_T)}{dz_T} \\ &+ \frac{(\kappa_2(P) - \kappa_2(N))}{2!}\frac{d^2n(z_T)}{dz_T^2} - \frac{(\kappa_3(P) - \kappa_3(N))}{3!}\frac{d^3n(z_T)}{dz_T^3} \\ &+ \frac{\left(\kappa_4(P) - \kappa_4(N) + 3(\kappa_2(P) - \kappa_2(N))^2\right)}{4!}\frac{d^4n(z_T)}{dz_T^4} + \varepsilon(z_T), \end{aligned} \tag{17}$$

where $\varepsilon(z_T)$ is an approximation error arising from the exclusion of higher order terms in the expansion and κ_i is the i^{th} cumulant of the associated distribution. This expression can be simplified by noting that both returns distributions are standardized to have zero mean

$(\kappa_1 = 0)$ and unit variance $(\kappa_2 = 1)$,

$$p(z_T) = n(z_T) - \frac{(\kappa_3(P) - \kappa_3(N))}{3!}\frac{d^3 n(z_T)}{dz_T^3} + \frac{(\kappa_4(P) - \kappa_4(N))}{4!}\frac{d^4 n(z_T)}{dz_T^4} + \varepsilon(z_T). \tag{18}$$

By using the properties of the normal distribution,

$$p(z_T) = n(z_T)\left[1 + \gamma_1 \frac{(z_T^3 - 3z_T)}{6} + \gamma_2 \frac{(z_T^4 - 6z_T^2 + 3)}{24}\right] + \varepsilon(z_T), \tag{19}$$

where

$$\gamma_1 = \frac{(\kappa_3(P) - \kappa_3(N))}{3!} \text{ and } \gamma_2 = \frac{(\kappa_4(P) - \kappa_4(N))}{4!} \tag{20}$$

are the unknown parameters which capture respectively skewness and kurtosis. The expression in (19) shows that the difference between the two densities, $p(z_T)$ and $n(z_T)$, is determined by the third and fourth moments. Substituting (19) into (8) and ignoring the approximation error, gives the approximate risk neutral probability distribution function

$$g(S_T|S_t) = |J|\, n(z_T)\left[1 + \gamma_1 \frac{(z_T^3 - 3z_T)}{6} + \gamma_2 \frac{(z_T^4 - 6z_T^2 + 3)}{24}\right], \tag{21}$$

where J is the Jacobian of the transformation from z_T to S_T given in (9) and (3) is used to substitute S_T for z_T. This expression shows that the density $g(S_T|S_t)$ is approximated by the lognormal distribution plus higher order terms which capture skewness and kurtosis. Using (21) in (2) to price options yields the Jarrow-Rudd option pricing model, augmented by time-varying volatility as defined in (4),

$$F(S_t) = BS + \gamma_1 Q_3 + \gamma_2 Q_4, \tag{22}$$

where BS is the Black-Scholes price defined in (16),

$$Q_3 = e^{-r\tau} \int_X^\infty (S_T - X)\,|J|\,\frac{(z_T^3 - 3z_T)}{6} dS_T \tag{23}$$

and

$$Q_4 = e^{-r\tau} \int_X^\infty (S_T - X)\,|J|\,\frac{(z_T^4 - 6z_T^2 + 3)}{24} dS_T, \tag{24}$$

with the substitution of S_T for z_T occurring via (3). In the case where volatility is assumed to be constant and $\beta_2 = 0$ in (4), (23) and (24) have analytical solutions as given in Jarrow and Rudd (1982). Otherwise, (23) and (24) need to be computed numerically.

The establishment of the relationship between the GST and Jarrow-Rudd models also highlights the risk neutral properties of the GST model. In particular, as the lognormal distribution corresponds to the risk neutral distribution for S_T, provided that the mean of the underlying return process is set equal to the risk free interest rate, and given that the GST distribution for S_T is constrained to have the same mean, it follows from the arguments in Jarrow and Rudd (1982) that the GST specification for $g(S_T|S_t)$ can be interpreted as the risk neutral probability distribution, at least in the local region around the lognormal distribution. In general, functions of the higher order moments in the risk neutral GST model, including variance, skewness and kurtosis, differ from the corresponding empirical functions. As the approach adopted in the paper is to specify the risk neutral distribution directly, the form of the adjustment in the higher order moments is taken as given.[6]

To highlight the ability of the Jarrow-Rudd model to approximate nonnormal distributions, the following experiment is performed, whereby the Jarrow-Rudd distribution is used to approximate the GST distribution. The results are presented in Figure 1 for various parameterizations of the GST distribution. The Jarrow-Rudd approximating distribution is given by (19) with the cumulants in (20) set equal to their respective values.

In Figure 1(a), the approximation is reasonable. In Figure 1(b), the Jarrow-Rudd approximating distribution does well in capturing the tails of the GST distribution, but over-estimates the peak. In Figure 1(c), the over-estimate of the peak of the GST distribution is more severe than it is in Figure 1(b), resulting in the approximating distribution over compensating in the approximation of the left tail. The approximation error is even more dramatic in 1(d), where the failure to approximate the peak of the true distribution correctly results in a spurious lobe in the left-hand tail of the approximating distribution.

One problem with implementing the Jarrow-Rudd approach is that the risk neutral probability distribution in (21) is not constrained to be positive over the support of the density. Interestingly, this problem appears to have been ignored in the literature so far. The problem is highlighted in (19) which shows that the returns distribution is a function of cubic and quartic polynomials, which can, in general, yield negative values. The problem of negativity can be expected to be more severe when the Jarrow-Rudd approximation does not model the true distribution accurately, causing the polynomial terms to over-adjust, especially in

[6] The links between higher order moments in the empirical and risk-neutral distributions can be identified numerically using the approach of Duan (1999). This procedure is based on sequential applications of the inverse cumulative distribution function technique.

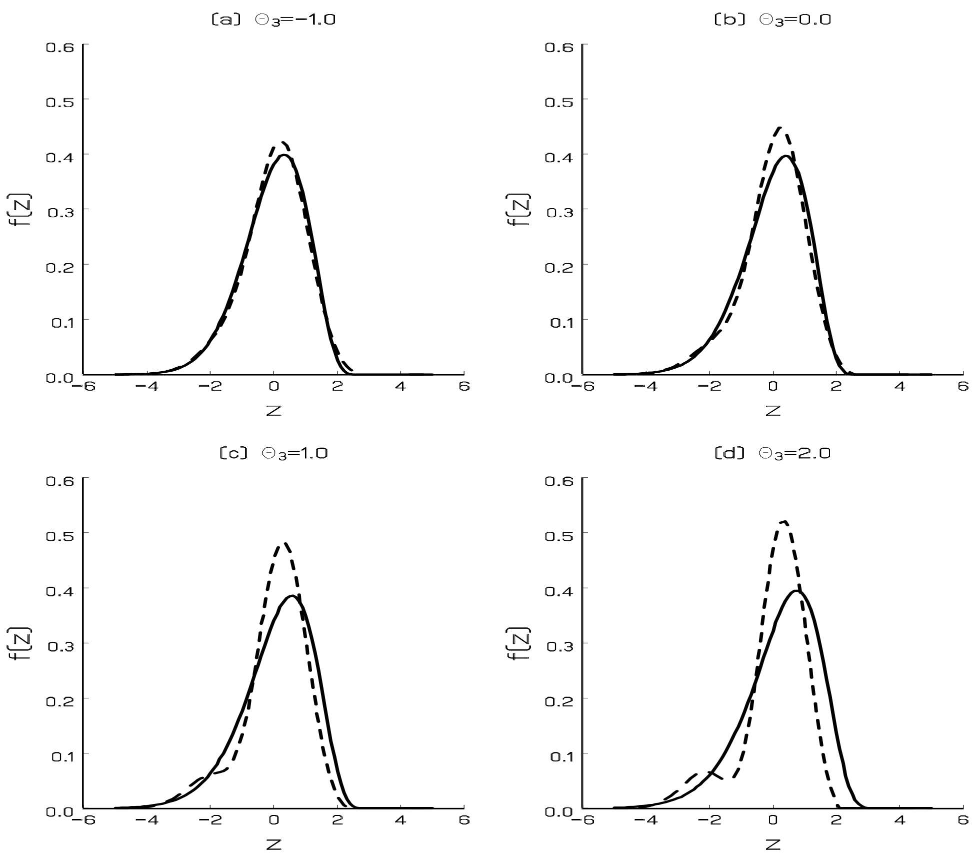

Figure 1: Approximation properties of the Jarrow-Rudd distribution: true distribution ——; approximating distribution - - - - - . The true distribution is a GST distribution with $\nu = 0.04$, $\theta_2 = -(1+\nu)/2$, $\theta_1 = -20$, $\theta_4 = -0.5$, $\theta_5 = 0$ and $\theta_6 = 0$.

the tails of the estimated distribution.[7]

To impose non-negativity on the underlying risk neutral probability distribution, an alternative specification is based on the semi-nonparametric density of Gallant and Tauchen (1989). Ignoring the approximation error, the standardized returns distribution in (19) is respecified as

$$p(z_T) = n(z_T)\left[1 + \lambda_1 \frac{(z_T^3 - 3z_T)}{6} + \lambda_2 \frac{(z_T^4 - 6z_T^2 + 3)}{24}\right]^2, \tag{25}$$

where the augmenting polynomial is now squared, forcing the probabilities to be greater than or equal to zero for all values of z_T. Using (25) in (2) yields the semi-nonparametric option pricing model which is compared with the GST model in the empirical application in Section 5. This model is referred to as the SNP option pricing model hereinafter.

3.3 Mixture of Lognormals

An alternative approach suggested by Melick and Thomas (1997) to capture departures from normal returns is based on a mixture of lognormal distributions. The option pricing model is

$$F(S_t) = \alpha BS(\sigma_1) + (1-\alpha) BS(\sigma_2), \tag{26}$$

where $BS(\sigma_i), i = 1, 2$, is the Black-Scholes price as defined in (16), and assuming constant volatility, σ_i, $i = 1, 2$. The parameter $0 \leq \alpha \leq 1$, is the mixing parameter which weights the two subordinate lognormal distributions.[8] In comparing alternative option price models in the empirical application below, for commensurability the mixture model is extended to allow for time-varying volatility, with σ_1 and σ_2 in (26) replaced respectively by

$$\begin{aligned} \sigma_{1,T|t} &= \left(\exp \beta_{1,1} + \beta_{1,2} \ln(S_T/S_t)\right) \\ \sigma_{2,T|t} &= \left(\exp \beta_{2,1} + \beta_{2,2} \ln(S_T/S_t)\right). \end{aligned} \tag{27}$$

[7] Negative probability estimates of the Jarrow-Rudd distribution were generated in the empirical application conducted in Section 4. These results are available from the authors upon request. For an earlier example of the problem of negative probabilities when using an Edgeworth expansion, see Kendall and Stuart (1969, p.160).

[8] In constrast to the specification in (26), in the lognormal mixture model proposed by Melick and Thomas (1997), risk-neutrality is not imposed on the underlying distribution.

4 Properties of the GST Option Pricing Model

4.1 Risk Neutral Distributional Shapes

Some examples of the risk neutral probability distribution $g(S_T|S_t)$ in (8) are given in Figures 2 and 3, for various parameterizations. The initial spot price is $S_t = 500$, for a 6 month option, $\tau = 6/12$, with a risk-free rate of interest of $r = 0.05$. In Figure 2 the effects of time-varying volatility on the risk-neutral probability distribution, assuming normal returns, are illustrated. The volatility specification is given by (4) with $\beta_1 = -2$, and a range of values of β_2 to control for the relative impact of expected returns over the life of the option on time-varying volatility. The risk neutral probability distribution of the Black-Scholes model is represented by the case $\beta_2 = 0$, in Figure 2. The results show that as β_2 increases the risk neutral probability distributions become more positively skewed.

In Figure 3, the effect of departures from normal returns on the shape of the risk neutral probability distribution is demonstrated. Three parameterizations of the GST-1 model are adopted, based on changes in ν, θ_1 and θ_4, with $\theta_3 = \theta_5 = \theta_6 = 0$. Setting $\nu = 4$ and $\theta_4 = -0.5$ causes the risk neutral probability distribution to become relatively more peaked than the lognormal distribution, with fatter tails. For $\theta_1 = 2$, the risk neutral probability distribution becomes even more positively skewed, whilst for $\theta_1 = -2$, the distribution exhibits relatively less positive skewness than the lognormal distribution.

4.2 Option Price Sensitivities

The effects of changes in the parameters of the GST-3 model on the option prices, is highlighted in Table 1. The option prices are computed for both one month contracts, $\tau = 1/12$, and three month contracts, $\tau = 3/12$, with strike prices of $X = 450, 500, 550$. The spot rate is $S_t = 500$ with a risk free rate of interest $r = 0.05$.

The Black-Scholes model is represented by the row labeled, Normal: $\beta_2 = 0$, that is, normal returns with constant volatility. The effects on the Black-Scholes price of time-varying volatility are highlighted in the next set of rows, with values of β_2 increasing from 0.1 to 0.4. Not surprisingly, option prices increase monotonically as the value of β_2 increases, reflecting that increases in risk caused by increases in volatility are priced at a premium.

The effects of changes in the parameters θ_3 and θ_5 in the returns distribution in (7) are highlighted by the rows labeled GST-3 in Table 1. Comparing the Black-Scholes and GST-3

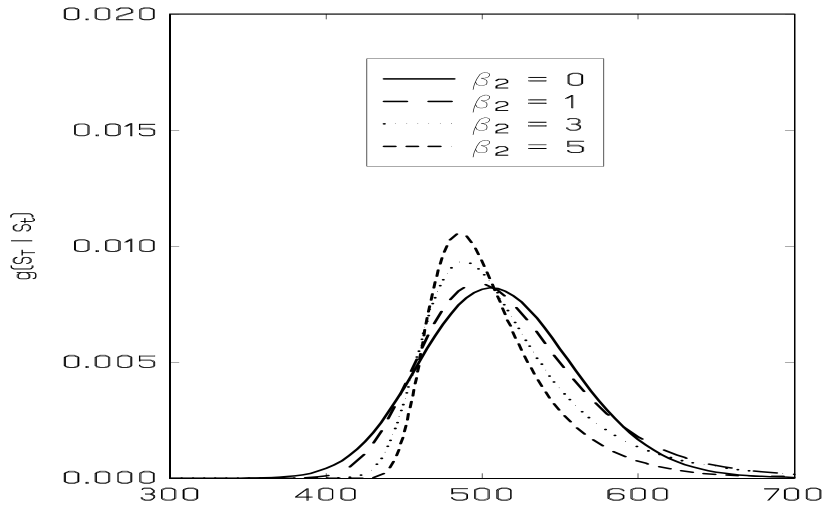

Figure 2: Risk neutral probability distributions for alternative volatility parameterisations, assuming normality in stock returns.

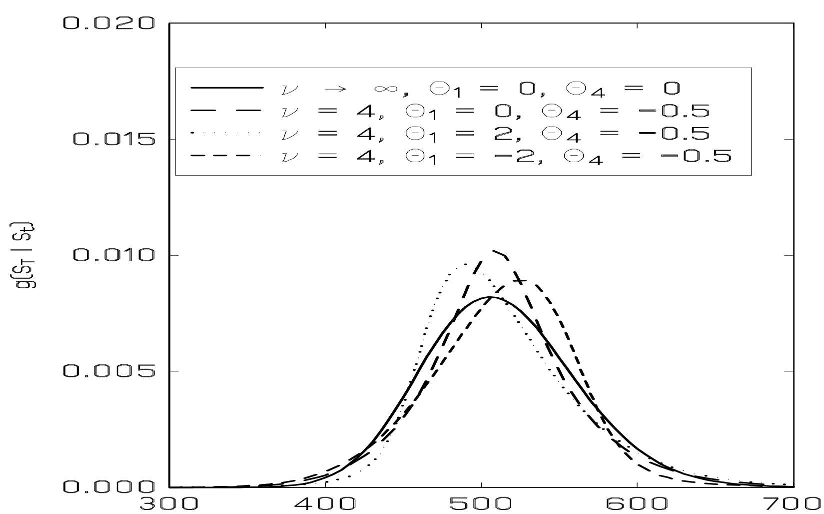

Figure 3: Risk neutral probability distributions for alternative distributional parameterisations of the GST-1 model in (10), with $\beta_2 = 0$, $\theta_2 = -(1+\nu)/2$ and $\theta_3 = \theta_5 = \theta_6 = 0$.

prices, it can be seen that the shorter-term GST-3 prices ($\tau = 1/12$) deviate relatively more from the corresponding Black-Scholes prices than do the longer-term prices ($\tau = 3/12$). For the one-month options, the results show that for the in-the-money options ($X = 450$), the Black-Scholes price, 51.88, is smaller than the GST-3 prices. For the at-the-money options ($X = 500$), the Black-Scholes price of 8.86 is slightly smaller than all of the GST-3 prices, except for the case where $\theta_3 = -2$ and $\theta_5 = -0.2$. For the out-the-money options ($X = 550$), the Black-Scholes price of 0.07 is less than the GST-3 prices for $\theta_3 = -2$, but equal to the GST-3 prices for $\theta_3 = -1$. Considering the results for the three-month options, the Black-Scholes prices are smaller than the GST-3 prices for $\theta_3 = -2$ for the in-the-money options and greater for $\theta_3 = -1$. The opposite situation occurs for the at-the-money-options. For the out-of-the money options, the Black-Scholes prices are less than the GST-3 prices for all parameterizations.

4.3 Volatility Skews

The relative differences in the Black-Scholes and GST-3 prices across moneyness groups in Table 1 demonstrate that the Black-Scholes implied volatilities are not constant, if options are priced according to the GST model. To highlight the relationship between volatility skews and misspecification of the returns distribution, the following experiments are performed. The experiments are based on a true volatility parameter value of $\sigma_{T|t} = \sigma = 0.15$ or 15%. The option prices are computed for a three month contract length, $\tau = 3/12$, based on a spot rate of $S = 500$, and strike prices ranging from $X = 400$ to $X = 600$, in steps of 1. The risk free rate of interest is $r = 0.05$.

Two experiments are conducted to highlight the relationship between the volatility skew and nonnormality in returns. In the first experiment, the effects on the implied volatility skew of changes in θ_5 in (7) are demonstrated. The values of the remaining parameters are specified in the caption for Figure 4. Equating this price with the Black-Scholes price gives the value of the implied volatility plotted in Figure 4.[9] The calculations are performed for $\theta_5 = -0.2, -0.1$. The results show that the volatility skew is reduced as the absolute value of θ_5 decreases. For deep in-the-money options ($X/S < 0.9$), there are large differences between the implied volatility and the true volatility of 0.15. This implies that Black-Scholes is seriously underpricing deep in-the-money options. For those contracts where the implied

[9] The GAUSS procedure *OPTMUM* is used to compute the implied volatilities.

Table 1:

Sensitivity of option prices to alternative conditional volatility parameterizations and distributional models
(For all models, $S_t = 500$, $r = 0.05, \beta_1 = -2$)

Distributional Model		$\tau = 1/12$ years			$\tau = 3/12$ years	
X :	450	500	550	450	500	550
Normal: $\beta_2 = 0.0$	51.88	8.86	0.07	56.12	16.74	1.88
Normal: $\beta_2 = 0.1$	51.96	8.90	0.07	56.31	16.88	1.98
Normal: $\beta_2 = 0.2$	52.03	8.95	0.08	56.51	17.02	2.07
Normal: $\beta_2 = 0.3$	52.11	8.99	0.09	56.71	17.17	2.18
Normal: $\beta_2 = 0.4$	52.19	9.03	0.10	56.92	17.32	2.28
GST-3$^{(a)}$: $\theta_3 = -2, \theta_5 = -0.2$	56.68	8.84	0.12	57.65	16.66	2.33
GST-3$^{(a)}$: $\theta_3 = -2, \theta_5 = -0.1$	54.29	8.87	0.11	56.90	16.73	2.26
GST-3$^{(a)}$: $\theta_3 = -1, \theta_5 = -0.2$	53.26	8.93	0.07	56.07	16.85	2.09
GST-3$^{(a)}$: $\theta_3 = -1, \theta_5 = -0.1$	52.97	8.92	0.07	56.04	16.84	2.06

(a) $\beta_2 = 0$, $\nu = 1, \theta_1 = 0, \theta_2 = -(1+\nu)/2, \theta_4 = -0.5, \theta_6 = -0.25$.

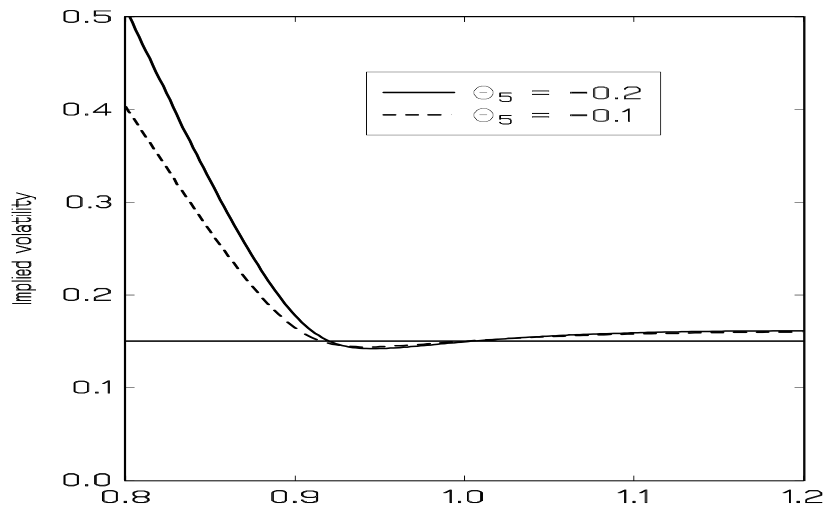

Figure 4: Volatility skews generated for $\tau = 0.25$ when returns are distributed as GST-3 with $\nu = 1$, $\theta_1 = 0$, $\theta_2 = -(1+\nu)/2$, $\theta_3 = -2$, $\theta_4 = -0.5$, $\theta_6 = -0.25$. The true volatility is $\sigma = 0.15$.

volatility is less than the true volatility of 0.15, Black-Scholes is overpricing. The differences between the implied and true volatilities are relatively small across the moneyness spectrum for $X/S \geq 0.9$.

The second volatility skew experiment is the same as the first, except that θ_3 is now allowed to vary. The values specified for the other parameters are given in the caption of Figure 5. The results presented in Figure 5 are qualitatively similar to those presented in Figure 4. In particular, there is strong evidence of a volatility skew, with large differences between the implied and true volatilities for deep in-the-money options. Increasing the value of θ_3 from -2.0 to -1.0 moderates the extent of the skew.

As an alternative experiment to identify the ability of the GST model to price options better than the model which does not allow for higher order moments, option prices are now simulated assuming that a volatility skew does indeed exist. The true price of a call option is equal to the Black-Scholes price, but with the volatility changing over moneyness. The idea is to see if the GST model can still recover the implied volatility skew, even when the underlying returns distribution is normal.[10] The true volatility skew is presented in Figure

[10] This experiment is based on the suggestion of an anonymous referee.

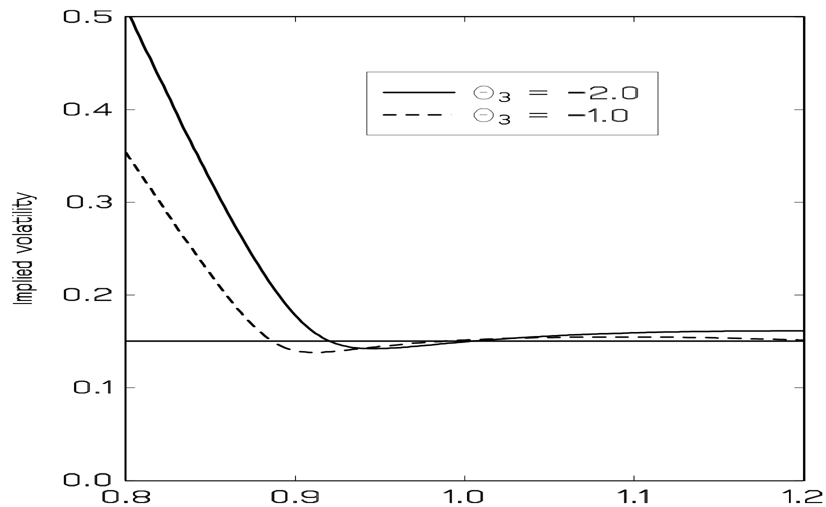

Figure 5: Volatility skews generated for $\tau = 0.25$ when returns are distributed as GST-3 with $\nu = 1$, $\theta_1 = 0$, $\theta_2 = -(1+\nu)/2$, $\theta_4 = -0.5$, $\theta_5 = -0.2$, $\theta_6 = -0.25$. The true volatility is $\sigma = 0.15$.

6. The volatility skew is constructed from call option data written on S&P500 stock options on April 4th and parameterized by using a quadratic regression equation in moneyness. This data is also used in the empirical section of the paper. The spot price is set at 500, with strike prices ranging from 450 to 550 in steps of 1, producing a total of 101 option contracts. Maturity is 3 months, $\tau = 3/12$, and the risk free rate of interest is set at $r = 0.05$. The simulated call option price associated with each contract is set equal to the Black-Scholes price, assuming that the true volatility is given by the volatility skew function. Both the normal and GST models are then estimated using the simulated data, which, in turn, are used to price the 101 option contracts.[11] Having generated the expected prices for the two models, the implied volatilities are then calculated by equating the expected price from each model with the Black-Scholes price.

The implied volatility smiles of the normal and GST models for the experiment are given in Figure 6. The normal model does poorly in recovering the volatility skew. In contrast the GST model does a very good job in recovering the volatility skew across the full moneyness

[11] Both the normal and GST models allow for a conditional volatility structure based on (4). With reference to (10), the GST-1 specification is used. The free parameters are estimated by maximum likelihood; see the discussion below about estimation methods.

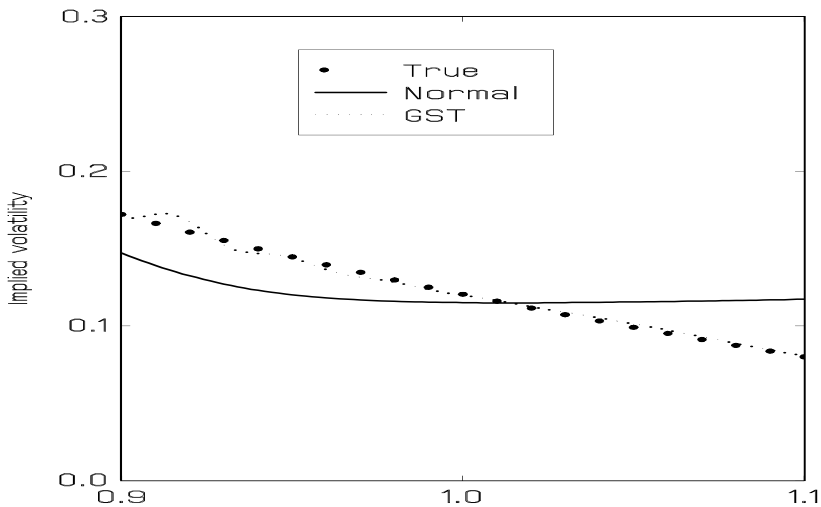

Figure 6: Simulation experiment to compare the ability of the GST-1 and normal option price models to recover a volatility skew.

spectrum.

5 Pricing S&P500 Stock Options

5.1 Data Description

The data set used in the empirical application consists of quotes on call options written on the S&P500 stock index, obtained from the Berkeley Options Database. The quotes relate to options traded in the month of April, 1995. Specifically, the alternative models are estimated using the midpoints of bid-ask quotes for April 4th, 11th and 18th respectively. Predictive and hedging performances are then assessed using data for the remaining days in April. Each sub-sample comprises approximately 40,000 prices, for option contracts extending over the full moneyness spectrum. Defining $S - X$ as the intrinsic value of the call option, options for which $S/X \in (0.97, 1.03)$ are categorized as at-the-money, those for which $S/X \leq 0.97$, as out-of-the-money, and those for which $S/X \geq 1.03$, as in-the-money. Maturity lengths range from approximately one to five months. Each record in the dataset comprises the bid-ask quote, the synchronously recorded spot price of the index, the time at which the quote was recorded, and the strike price.

As dividends are paid on the S&P500 index, the current spot price, S_t, in (1) and in all subsequent formulae is replaced by the dividend-exclusive spot price, $S_t e^{-d\tau}$, where d is the average rate of dividends paid on the S&P500 index over 1995. This rate is used as a proxy for the rate of dividend payment made over the life of each option. Daily dividend data for 1995, used to construct d, were obtained from Standard and Poors. Only observations for which the average of the bid and ask prices exceeds the lower bound of

$$LB = \max\{0, S_t e^{-d\tau} - e^{-r\tau} X\},$$

and which are recorded between 9.00am and 3.00pm are included in the sample. The first restriction serves to exclude prices which fail to satisfy the no-arbitrage lower bound, whilst the second restriction seeks to minimize the problem of nonsynchroneity between the spot and option prices.

The interest rate, r, is the three month bill rate observed on that day, with interest rate data obtained from Datastream. Tables 2, 3 and 4 summarize the main characteristics of the datasets used in the estimation.

5.2 Model Estimation

Define the theoretical price of the j^{th} option contract at time t, as

$$F_{j,t} = F\left(S_t, X_j, \tau_j, r; \Omega\right), \tag{28}$$

where Ω is the vector of parameters which characterize the returns distribution and the volatility specification. In the case of the Black-Scholes option pricing model, for example, $\Omega = \{\beta_1\}$. The relationship between $C_{j,t}$, the market price of the j^{th} option contract at time t, and $F_{j,t}$, is given by

$$C_{j,t} = F_{j,t} + e_{j,t}, \tag{29}$$

where $e_{j,t}$ represents the pricing error. This pricing error can be viewed as capturing the deviation which occurs between an observed market option price, $C_{j,t}$, and any version of the theoretical price, $F_{j,t}$. For example, a deviation may occur due to market-related factors, such as the non-synchronous recording of spot and option prices. The error can also be viewed as encompassing the fact that any parametric model is only ever an approximation of the process by which market participants set an option price. Following the approach of Engle

Table 2:

S&P500 Option Price Dataset: April 4, 1995.

Variable		All Contracts	Moneyness (S/X) < 0.97	Moneyness (S/X) 0.97 – 1.03	Moneyness (S/X) > 1.03
Call Price:	$\overline{X}$	\$60.87	\$8.60	\$19.90	\$66.36
	SD	\$31.17	\$2.84	\$5.82	\$28.83
	Min	\$1.04	\$1.04	\$3.38	\$21.63
	Max	\$156.25	\$10.63	\$29.88	\$156.25
	$Number$	43584	511	4519	38554
Maturity (No. of Prices)	May	11600	45	121	11434
	June	16748	39	2047	14662
	Sept.	15236	427	2351	12458
Strike Price:	$\overline{X}$	\$448.65			
	SD	\$34.02			
	Min	\$350.00			
	Max	\$550.00			
	$Number$	30			
Spot: (S&P 500 Index)	$\overline{X}$	\$503.33			
	SD	\$0.56			
	Min	\$502.38			
	Max	\$504.56			

Table 3:

S&P500 Option Price Dataset: April 11, 1995.

Variable		All Contracts	Moneyness (S/X) < 0.97	$0.97 - 1.03$	> 1.03
Call Price:	$\overline{X}$	\$50.00	\$6.31	\$17.13	\$57.01
	SD	\$27.15	\$1.98	\$5.33	\$24.50
	Min	\$0.95	\$0.95	\$1.94	\$19.13
	Max	\$157.63	\$10.07	\$29.00	\$157.63
	$Number$	43509	438	7088	35983
Maturity:	May	9609	19	1706	7884
(No. of Prices)	June	15254	20	1868	13366
	Sept.	18646	399	3514	14733
Strike Price:	$\overline{X}$	\$463.37			
	SD	\$29.96			
	Min	\$350.00			
	Max	\$600.00			
	$Number$	31			
Spot:	$\overline{X}$	\$506.19			
(S&P 500 Index)	SD	\$0.72			
	Min	\$502.29			
	Max	\$508.42			

Table 4:

S&P500 Option Price Dataset: April 18, 1995.

Variable		All Contracts	Moneyness (S/X)		
			< 0.97	$0.97 - 1.03$	> 1.03
Call Price:	$\overline{X}$	\$46.89	\$5.13	\$17.38	\$53.73
	SD	\$25.84	\$2.45	\$5.19	\$23.48
	Min	\$0.60	\$0.60	\$1.10	\$17.38
	Max	\$156.88	\$8.63	\$29.50	\$156.88
	$Number$	38099	670	6274	31155
Maturity: (No. of Prices)	May	7984	25	1297	6662
	June	12579	27	1699	10853
	Sept.	17536	618	3278	13640
Strike Price:	$\overline{X}$	\$465.73			
	SD	\$28.67			
	Min	\$350.00			
	Max	\$550.00			
	$Number$	30			
Spot: (S&P 500 Index)	$\overline{X}$	\$505.65			
	SD	\$0.63			
	Min	\$504.12			
	Max	\$506.71			

and Mustafa (1992), Sabbatini and Linton (1998) and Jacquier and Jarrow (2000), the pricing error is assumed to be a normal random variable with zero mean; see also the discussion in Clement, Gourieroux and Monfort (2000). To allow for the variance of the pricing error to vary across option contracts of different degrees of moneyness, $e_{j,t}$ is specified as having a variance $\omega_{j,t}^2$, defined by

$$\omega_{j,t}^2 = \phi_1 + \phi_2 (S_t / X_j), \tag{30}$$

where ϕ_1 and ϕ_2 are unknown parameters.

Letting N represent the number of observations in a pooled data set of time series and cross-sectional prices of option contracts, the logarithm of the likelihood function is given by

$$\ln L = -\frac{1}{2} \sum_{j,t}^{N} \ln\left(2\pi\omega_{j,t}^2\right) - \frac{1}{2} \sum_{j,t}^{N} \left(\frac{C_{j,t} - F_{j,t}}{\omega_{j,t}}\right)^2 . \tag{31}$$

This function is maximized with respect to ϕ_1, ϕ_2 and Ω, using the GAUSS procedure MAXLIK. The numerical integration procedure for computing the theoretical option price $F_{j,t}$ for the various models is based on the GAUSS procedure INTQUAD1. All integrals are transformed to the $(-1, 1)$ range. As a test of the accuracy of the integration procedure, both numerical and analytical formulae for the Black-Scholes model were used. Both procedures generated the same parameter estimates to at least four decimal points.[12]

5.3 Performance Evaluation

The performance of the alternative pricing models is now investigated. Five procedures are used to assess the performance of the models. The first consists of conducting standard tests of significance on the parameter estimates. The second concentrates on comparing the relative size of mispricing errors of each model. The third focuses on forecasting properties, whilst the fourth procedure compares the relative size of hedging errors from each model. The last procedure examines the ability of the competing models to correct for volatility skews.

5.3.1 Statistical Tests

Tables 5 to 7 contain the parameter estimates of the GST and normal option price models for the three respective days investigated, with standard errors based on the inverse of the

[12] Computation time for all models estimated is rapid. For example, in the case of the most highly parameterized model (GST-3), the GAUSS program takes 42 iterations to converge. The total time taken to achieve convergence is just over one hour on a Pentium 4, 200 Mhz.

Hessian given in parentheses. The models consist of the three GST models as defined in (10), and the normal distribution. All models use the volatility specification in (4).

The parameter estimates of β_2 reported in Tables 5 to 7 are statistically significant for all models across all three days, thereby providing strong evidence of time-varying volatility. In particular, the significant estimate of β_2 in the normal model across all three days, represents an immediate rejection of the Black-Scholes model, which assumes normal returns and constant volatility.

The estimates of the distributional parameters provide further rejection of the Black-Scholes model as they highlight strong evidence of nonnormality. For example, there is strong evidence of negative skewness across all three days and across all three GST models investigated. In the case of the GST-2 and GST-1 models, negative skewness is identified by the negative estimates of θ_3 and θ_1 respectively. In the case of the more general model, GST-3, skewness is jointly modelled by the two parameters θ_3 and θ_5, which capture odd moment behavior. To identify the sign of skewness for this model, the third moment of the returns distribution is calculated as $\int z_T^3 p(z_T)\, dz_T$, which shows that the GST-3 distribution is also negatively skewed across all three days.[13]

Measurement of kurtosis via the calculation of $\int z_T^4 p(z_T)\, dz_T$ indicates that the GST-1 and GST-2 models exhibit significant excess kurtosis. For the GST-3 model, the inclusion of the higher order term, $(\mu_w + \sigma_w z_T)^4$, in the exponential in (7), with a value of -0.25 for the associated parameter θ_6, serves to reduce kurtosis to slightly less than the value of 3 associated with the normal distribution, for all three days. For this model then, the more significant departure from normality occurs in the form of skewness.[14]

A comparison of the point estimates for all three GST models, across the three days investigated, shows that the signs are consistent and the magnitudes are similar, especially in the case of the point estimates based on April 11th and 18th options. This result is also supported by the estimates of skewness and kurtosis for each GST model across the three

[13] For example, the third moment around the mean of the returns distribution is computed as -0.235 using data on the 4th of April. Performing a test of skewness under the null hypothesis of normality leads to a strong rejection of the null using conventional significance levels. Similar results occur for the other two days investigated.

[14] Given the findings of significant departures from normality and constant volatility, the rules of Ait Sahalia (2002) are used to determine if the underlying continuous time process is a diffusion. The criterion for a diffusion is $\partial^2 \ln (g(S_T|S_t)) / \partial S_T \partial S_t > 0$, $\forall S_T, S_t$. Using data for April 4th, for example, negative values of the criterion are found. Using the delta method to compute the standard errors of the criterion function, these negative values are found to be statistically significant.

days.[15]

For comparison, the parameter estimates of the SNP and mixture of lognormals option pricing models are reported in Tables 8 and 9 respectively. Both models show significant evidence of time-varying volatility across all three days. Both models also provide evidence of departures from normality in returns, as indicated by the significance of all relevant parameters.

The estimates of ϕ_2 for all five models reported in Tables 5 to 9 and across all three days investigated, show a significant relationship between the pricing error variance and the moneyness of the option contract.[16]

5.3.2 Mispricing

An overall measure of the pricing error is given by

$$s^2 = \frac{\sum_{j,t} \left(C_{j,t} - F_{j,t}\right)^2}{N}, \tag{32}$$

where $C_{j,t}$ and $F_{j,t}$ are defined above.[17] Estimates of the residual variance for the most general of the GST models, GST-3, along with the SNP, mixture and normal models, across the three days, are given Table 10. To allow for differences in parameter dimensionality across the models when comparing mispricing properties, the AIC and SIC statistics are also presented. A comparison of all measures of fit across the models shows that the GST-3 model yields reductions in the amount of mispricing relative to the mixture and normal models, on all three days. It also exhibits less mispricing than the SNP model on April 4th and April 18th, producing slightly higher average mispricing than the SNP model on April 11th. Overall, the SNP model is the next best performer, followed by the mixture model. The normal model yields increases in mispricing over the GST-3 model of nearly 7000% on April 4th and between approximately 120% and 500% on the other two days. These statistics provide further strong evidence that there are large gains to be made from modelling the nonnormalities in stock returns and that the GST-3 model does a better job

[15] A formal test of parameter constancy across the three days, for each model, is conducted by comparing the log of the likelihood from pooling the data with the sum of the three log likelihoods from not pooling. This test leads to rejection of parameter constancy for each model. These results are not reported in the paper.

[16] The single exception to this finding is the insignificant estimate of ϕ_2 for the GST-3 model on April 18th.

[17] Whilst equation (32) is based on an $L2$ norm, other norms could be used to measure the extent of the pricing error.

overall in capturing the impact of these distributional features on option prices than do the alternative nonnormal specifications.[18]

5.3.3 Forecasting

The forecasting performance of the GST-3, SNP, mixture and normal pricing models are compared in Table 11. For each of the three days for which parameter estimates are produced (4th, 11th and 18th of April), forecasts for the remaining days of the corresponding week are computed. For example, the parameter estimates based on the 4th of April data are used to compute option prices on the 5th, 6th and 7th of April. The forecast error is given by

$$fe_{j,t} = C_{j,t} - F_{j,t|\text{April 4th}}, \tag{33}$$

where $C_{j,t}$ are call option prices recorded on either the 5th, 6th or 7th of April, and

$$F_{j,t|\text{April 4th}} = F\left(S_t, X_j, \tau_t, r; \Omega_{\text{April 4th}}\right), \tag{34}$$

with $\Omega_{\text{April 4th}}$ signifying parameter estimates based on April 4th data. The parameter estimates based on the 11th of April data are used to compute option prices for the 12th and 13th of April, and the parameter estimates based on the 18th of April data used to compute option prices for the 19th to the 21st of April, where the corresponding forecast errors are defined as in (33).

The forecasting performance of the models is assessed using two statistics. The first is the Diebold and Mariano (1995) (DM) statistic, which is used to test whether the differences between the pricing errors of any particular model and the normal model are significant. The statistic is calculated as

$$DM = \frac{\frac{1}{N}\sum_{j,t} d_{j,t}}{\sqrt{\frac{1}{N}\sum_{j,t} d_{j,t}^2}}, \tag{35}$$

where $d_{j,t}$ is the difference between the forecast errors of any of the models based on a nonnormal distribution and the model based on the normal distribution for the jth option contract at time t. Under the null hypothesis that there is no difference in the forecast errors of the two models, DM is asymptotically distributed as $N(0,1)$. The results in Table 11 show that in all but 2 of the 24 cases, the differences are significant.

[18] An alternative measure of mispricing, along the lines of that used in Corrado and Su (1997), would be based on the proportion of theoretical prices, for each option price model, which fall within the observed bid-ask spread associated with the option contract.

The second statistic used to evaluate the forecasting properties of the models is the Root Mean Squared Error (RMSE). This statistic assesses whether the nonnormal specifications produce errors which are smaller in magnitude, on average, than those produced by the normal specification. In the case of forecasts based on the April 4th data, the RMSE is computed as

$$RMSE = \sqrt{\frac{\sum_{j,t} fe_{j,t}^2}{N}}, \tag{36}$$

where $fe_{j,t}$ is as defined in (33). The results in Table 11 show that overall the GST-3 pricing model yields the smallest RMSE, followed by the SNP, mixture and normal models. In the three cases where the $RMSE$ of the SNP model is smaller than that of the GST-3 model, the differences are only marginal.

5.3.4 Hedging Errors

An important aim of risk management is to construct risk-free portfolios. The ability to achieve this aim is not only a function of the frequency with which a portfolio is rebalanced, but also a function of the accuracy with which the assumed model explains the data. To examine this latter property, the competing models are now used to construct portfolios based on delta hedges.

Consider forming a portfolio that is short in the call option. Normalizing the portfolio on a single call option contract, the size of the investment, $I_{j,t}$, required to set up the portfolio, $P_{j,t}$, using a delta hedge is

$$P_{j,t} = I_{j,t} = \Delta_{j,t} S_t - C_{j,t}, \tag{37}$$

where S_t is the spot price at the time the portfolio is constructed, $C_{j,t}$ is the call price on the j^{th} option contract, and $\Delta_{j,t}$ represents the proportion of stocks purchased to delta hedge the portfolio,

$$\Delta_{j,t} = \frac{dC_{j,t}}{dS_t}. \tag{38}$$

The value of the portfolio at the start of the next day, based on the proportion of stocks purchased in the previous day, is

$$P_{j,t+1} = \Delta_{j,t} S_{t+1} - C_{j,t+1}. \tag{39}$$

The value of this portfolio can be compared to investing the amount $I_{j,t}$ in (37) at the risk free rate of interest r for one day. The value of this investment in period $t+1$, is

$$I_{j,t+1} = I_{j,t} \exp\left(r/365\right) = \left(\Delta_{j,t} S_t - C_{j,t}\right) \exp\left(r/365\right). \tag{40}$$

The difference between (39) and (40) yields the one day ahead hedging error; see Bakshi, Cao and Chen (1997),

$$\begin{aligned} H_{j,t+1} &= P_{j,t+1} - I_{j,t+1} \\ &= \Delta_{j,t}\left(S_{t+1} - S_t \exp\left(r/365\right)\right) - \left(C_{j,t+1} - C_{j,t} \exp\left(r/365\right)\right). \end{aligned} \tag{41}$$

The hedging error for k days ahead is calculated as

$$H_{j,t+k} = \Delta_{j,t}\left(S_{t+k} - S_t \exp\left(rk/365\right)\right) - \left(C_{j,t+k} - C_{j,t} \exp\left(rk/365\right)\right), \tag{42}$$

which is the measure used to compare the hedging performances of the competing models.[19]

The results of the hedging error experiments for the various option price models across the three sample periods are contained in Tables 12 to 14 respectively. These results give the incremental value over investing in a risk free asset from constructing a portfolio on the 4th, 11th and 18th of April which is not rebalanced over the time horizons given. The total number of unique contracts that have matching contracts across the time horizon are 574 for April 4th results, 563 for April 11th results and 377 for April 18th results.[20] All values are expressed in dollars whereby a value of $+X$ $(-X)$ means that the portfolio earns $\$X$ more than (less than) would be earned from investing the money at the risk free rate of interest over the pertinent forecast period. The size of the hedging errors are broken-down into moneyness classes, S_t/X_j, as well as being reported for the total class. For comparison, the average values of the investment, $I_{j,t}$ in (37), across all contracts for each model are also presented in Tables 12 to 14. In calculating $\Delta_{j,t}$ in (38) for each of the models, a numerical differentiation procedure is used.

Concentrating on the hedging errors associated with all contracts in Tables 12 to 14, across all hedging periods, the GST-3 models produces smaller absolute hedging errors than the normal model, on average, in 6 of the 8 cases.[21] In comparing across nonnormal models, especially the GST-3 and SNP models, there is very little difference in average hedging performance. To assess the significance of the difference between the hedging errors of the

[19] Although volatility is stochastic, as discussed earlier, is it still a function of the single error process which drives the mean of the underlying spot price. This implies that a vega hedge is not required, which is in contrast with the class of volatility models used by Heston (1993), for example. However, in calculating the delta in (38), as S_t enters the volatility specification in (4), the hedging portfolio is affected by the time-varying nature of the volatility.

[20] In computing the hedging errors, the spot rates are those time-stamped with the call option. Hence, the spot rates will vary slightly over the day for different contracts written on the same day.

[21] Preliminary tests (not reported) that the average magnitude of the hedging errors associated with each model are zero, leads to strong rejection.

nonnormal and normal models, Tables 12 to 14 also include the DM statistic from (35), where the forecast errors are replaced by the hedging errors in (42). These results show that in the great majority of cases, the hedging errors associated with the nonnormal specifications are significantly different from the errors associated with the normal model.

One feature of the hedging errors in Tables 12 to 14 is that the hedging errors for all models increase in magnitude as the hedging horizon increases. This may not only reflect the lack of rebalancing of the portfolio, but also the fact that the delta hedge is not necessarily the appropriate risk management instrument to use when the returns process underlying the option contract is nonnormal. An interesting future research project would be to devise hedging strategies to manage higher order moments in the underlying returns.

5.3.5 Volatility Skew Corrections

As a final performance measure, the ability of the GST model to correct for volatility skews is examined. The results are given in Figure 7 which compares the implied volatility skew for the Black-Scholes model with the implied volatility skew obtained for the GST model, using options prices on April 4th 1995, written on May contracts.[22] In computing the option price, $F_{j,t}$, the point estimates of the GST distribution are based on the GST-3 model given in Table 5. The implied volatility parameter is computed by solving

$$C_{j,t} = F_{j,t}(\sigma),$$

for each contract assuming volatility over the life of the contract is fixed, that is, $\sigma_{T|t} = \sigma$. The calculations are performed over the full range of strike prices.[23] To generate a smooth implied volatility surface, the implied volatility estimates presented in Figure 7 are the predictions from regressing the implied volatility values on a constant and a quadratic polynomial in moneyness.

For comparability with Figures 4 and 5, the volatility smiles presented in Figure 7 are plotted against the inverse of moneyness, X/S. The results show the volatility skew associated with the Black-Scholes model, with implied volatility values of just under 40% for deep in-the-money contracts, and about 10% for deep out-of-the-money contracts. The

[22] Similar qualitative results are obtained for the June and September contracts, as well as from using data on April 11th and 18th, 1995. To save space these results are not presented.

[23] Contracts with the same strike price X, but different moneyness as a result of differences in the spot price S_t, over the day, are all included. This yields a total sample size of over $11,000$ contracts to compute the implied volatility functions for April 4th.

Table 5:

Maximum likelihood estimates of option price models for the 4th of April 1995: standard errors in brackets, $N = 43584$.

Parameter	GST-3	GST-2	GST-1	Normal
β_1	-2.107	-2.033	-2.001	-2.003
	(0.001)	(0.001)	(0.001)	(0.001)
β_2	0.208	0.189	0.163	0.178
	(0.002)	(0.001)	(0.001)	(0.002)
$\gamma = \sqrt{\nu}$	0.655	0.575	0.569	n.a.[(a)]
	(0.004)	(0.002)	(0.002)	
θ_1	0.000	0.000	-0.515	0.000
			(0.002)	
θ_2	-0.715	-0.665	-0.662	0.000
	(0.002)	(0.001)	(0.001)	
θ_3	0.708	-0.373	0.000	0.000
	(0.005)	(0.001)		
θ_4	-1.072	-0.500	-0.500	-0.500
	(0.008)			
θ_5	-1.034	0.000	0.000	0.000
	(0.002)			
θ_6	-0.250	0.000	0.000	0.000
ϕ_1	-3.687	-6.387	-7.197	-7.820
	(0.046)	(0.048)	(0.046)	(0.051)
ϕ_2	2.538	6.111	7.097	8.038
	(0.051)	(0.054)	(0.051)	(0.057)
Av. log-likelihood[(b)]	-0.011	-0.517	-0.592	-0.814

(a) n.a. = not applicable.

(b) The average log-likelihood is lnL/N, where L is likelihood and N is the sample size.

Table 6:

Maximum likelihood estimates of option price models for the 11th of April 1995: standard errors in brackets, $N = 43509$.

Parameter	GST-3	GST-2	GST-1	Normal
β_1	-2.127	-2.054	-2.019	-2.108
	(0.001)	(0.001)	(0.001)	(0.001)
β_2	0.418	0.314	0.261	0.420
	(0.004)	(0.004)	(0.003)	(0.001)
$\gamma = \sqrt{\nu}$	0.592	0.355	0.338	n.a.[a]
	(0.007)	(0.001)	(0.001)	
θ_1	0.000	0.000	-0.292	0.000
			(0.002)	
θ_2	-0.675	-0.563	-0.557	0.000
	(0.004)	(0.001)	(0.001)	
θ_3	1.031	-0.314	0.000	0.000
	(0.012)	(0.003)		
θ_4	-0.338	-0.500	-0.500	-0.500
	(0.017)			
θ_5	-0.848	0.000	0.000	0.000
	(0.005)			
θ_6	-0.250	0.000	0.000	0.000
ϕ_1	0.251	-3.111	-4.357	-3.601
	(0.062)	(0.073)	(0.056)	(0.033)
ϕ_2	-1.426	2.714	4.156	3.638
	(0.067)	(0.079)	(0.061)	(0.036)
Av. log-likelihood[b]	-0.355	-0.810	-0.894	-1.172

(a) n.a. = not applicable.

(b) The average log-likelihood is lnL/N, where L is likelihood and N is the sample size.

Table 7:

Maximum likelihood estimates of option price models for the 18th of April 1995: standard errors in brackets, $N = 38099$.

Parameter	GST-3	GST-2	GST-1	Normal
β_1	-2.137 (0.009)	-2.069 (0.001)	-2.042 (0.001)	-2.120 (0.002)
β_2	0.350 (0.041)	0.270 (0.003)	0.243 (0.003)	0.421 (0.001)
$\gamma = \sqrt{\nu}$	0.592 (0.012)	0.444 (0.002)	0.436 (0.003)	n.a.[(a)]
θ_1	0.000	0.000	-0.484 (0.003)	0.000
θ_2	-0.675 (0.002)	-0.599 (0.001)	-0.595 (0.001)	0.000
θ_3	0.917 (0.021)	-0.370 (0.002)	0.000	0.000
θ_4	-0.344 (0.070)	-0.500	-0.500	-0.500
θ_5	-0.796 (0.019)	0.000	0.000	0.000
θ_6	-0.250	0.000	0.000	0.000
ϕ_1	-1.204 (0.850)	-4.245 (0.064)	-5.047 (0.062)	-4.128 (0.038)
ϕ_2	0.295 (0.917)	3.910 (0.069)	4.860 (0.067)	4.089 (0.041)
Av. log-likelihood[(b)]	-0.488	-0.800	-0.879	-1.083

(a) n.a. = not applicable.

(b) The average log-likelihood is lnL/N, where L is likelihood and N is the sample size.

Table 8:

Maximum likelihood estimates of the SNP
price model for various dates: standard errors in brackets.

Parameter	4th of April	11th of April	18th of April
β_1	-2.109 (0.001)	-2.219 (0.001)	-2.225 (0.002)
β_2	-0.185 (0.004)	0.436 (0.001)	0.434 (0.001)
λ_1	-0.386 (0.002)	-0.626 (0.002)	-0.576 (0.002)
λ_2	0.133 (0.003)	-0.048 (0.001)	-0.088 (0.001)
ϕ_1	-5.199 (0.066)	-0.238 (0.117)	0.293 (0.166)
ϕ_2	4.426 (0.073)	-0.965 (0.124)	-1.296 (0.172)
Av. log-likelihood$^{(a)}$	-0.193	-0.292	-0.511

(a) The average log-likelihood is lnL/N, where L is likelihood and N is the sample size.

Table 9:

Maximum likelihood estimates of the mixture of lognormal option price model for various dates: standard errors in brackets.

Parameter	4th of April	11th of April	18th of April
$\beta_{1,1}$	-1.415 (0.003)	-1.426 (0.003)	-1.429 (0.005)
$\beta_{1,2}$	0.149 (0.004)	-0.142 (0.004)	-0.355 (0.009)
$\beta_{2,1}$	-2.616 (0.001)	-2.708 (0.009)	-2.495 (0.006)
$\beta_{2,2}$	-0.301 (0.017)	0.513 (0.001)	0.488 (0.001)
α	0.274 (0.001)	0.284 (0.002)	0.228 (0.003)
ϕ_1	-6.853 (0.051)	-7.885 (0.070)	-9.248 (0.075)
ϕ_2	6.278 (0.057)	7.476 (0.076)	9.043 (0.081)
Av. log-likelihood$^{(a)}$	-0.201	-0.426	-0.557

(a) The average log-likelihood is lnL/N, where L is likelihood and N is the sample size.

Table 10:

Estimates of mispricing of alternative models across selected days.

Day	Statistic	GST-3	SNP	Mixture	Normal
4th of April	$s^{2(a)}$	0.065	0.114	0.180	0.839
	AIC$^{(b)}$	0.022	0.386	0.403	1.628
	SIC$^{(c)}$	0.023	0.387	0.404	1.629
11th of April	s^2	0.121	0.106	0.325	0.769
	AIC	0.710	0.583	0.853	2.343
	SIC	0.711	0.585	0.854	2.344
18th of April	s^2	0.156	0.168	0.536	0.671
	AIC	0.977	1.021	1.114	2.166
	SIC	0.978	1.023	1.116	2.167

(a) Based on equation (32).

(b) AIC = -2lnL/N+2k/N, where L is the likelihood, N is the sample size and k is the number of estimated parameters.

(c) SIC = -2lnL/N+ln(N)k/N, where L is the likelihood, N is the sample size and k is the number of estimated parameters.

Table 11:

Forecasting performance of alternative option price models across various days in April 1995; DM is the Diebold-Mariano statistic given by (35) and RMSE is the root mean square error defined in (36). (** indicates significance at the 5% level.)

Forecast Day	Statistic	GST-3	SNP	Mixture	Normal
			4th of April		
5th	DM	13.597**	2.002**	6.658**	n.a.[(a)]
	RMSE	0.327	0.414	0.516	0.929
6th	DM	20.427**	7.344**	14.093**	n.a.
	RMSE	0.331	0.414	0.532	0.979
7th	DM	0.580	-16.285**	-4.870**	n.a.
	RMSE	0.370	0.443	0.488	0.848
			11th of April		
12th	DM	-26.252**	-6.271**	-38.652**	n.a.
	RMSE	0.346	0.332	0.573	0.872
13th	DM	-37.600**	-20.756**	-53.934**	n.a.
	RMSE	0.506	0.501	0.698	0.900
			18th of April		
19th	DM	2.009**	37.007**	-37.298**	n.a.
	RMSE	0.419	0.534	0.789	0.880
20th	DM	-15.603**	0.772	-59.597**	n.a.
	RMSE	0.406	0.420	0.720	0.809
21st	DM	-30.250**	-29.778**	-68.538**	n.a.
	RMSE	0.968	0.952	1.092	1.112

(a) n.a. = not applicable.

Table 12:

Hedging performance of alternative option price models constructed on the 4th of April 1995: Average over contracts of excess profits expressed in dollars relative to a risk-free investment. (** indicates significance at the 5% level.)

Day		GST-3	SNP	Mixture	Normal
4th	Investment (I)	406.181	406.018	404.992	400.637
	Moneyness (S/X)				
5th	<0.97	-0.326	-0.342	-0.456	-0.396
	0.97 - 1.00	-0.205	-0.237	-0.281	-0.325
	1.00 - 1.03	0.151	0.122	0.113	0.031
	>1.03	0.164	0.166	0.166	0.160
	All contracts[(a)]	0.153	0.153	0.150	0.139
	DM[(b)]	6.307**	8.376**	5.553**	n.a.
6th	<0.97	-0.015	-0.037	-0.217	-0.118
	0.97 - 1.00	6.376	6.301	6.173	6.087
	1.00 - 1.03	1.510	1.455	1.428	1.282
	>1.03	-3.057	-3.058	-3.057	-3.086
	All contracts[(a)]	-2.606	-2.613	-2.615	-2.653
	DM	7.313**	8.213**	6.564**	n.a.
7th	<0.97	14.131	14.128	14.123	14.129
	0.97 - 1.00	13.814	13.791	13.786	13.727
	1.00 - 1.03	5.899	5.883	5.890	5.824
	>1.03	-5.930	-5.934	-5.932	-5.963
	All contracts[(a)]	-4.932	-4.937	-4.935	-4.968
	DM	6.303**	7.197**	6.331**	n.a.

n.a.= not applicable.

(a) These average hedging errors are all significantly different from zero.

(b) Diebold-Mariano statistic computed for all contracts.

Table 13:

Hedging performance of alternative option price models constructed on the 11th of April 1995: Average over contracts of excess profits expressed in dollars relative to a risk-free investment. (** indicates significance at the 5% level.)

Day		GST-3	SNP	Mixture	Normal
11th	Investment (I)	409.833	409.427	408.536	408.610
	Moneyness (S/X)				
12th	$<$0.97	-0.197	-0.197	-0.194	-0.195
	0.97 - 1.00	0.013	0.013	0.011	0.026
	1.00 - 1.03	-0.069	-0.071	-0.068	-0.038
	$>$1.03	-0.036	-0.036	-0.037	-0.039
	All contracts$^{(a)}$	-0.039	-0.039	-0.039	-0.037
	DM$^{(b)}$	-1.165	-1.270	-1.717	n.a.
13th	$<$0.97	-	-	-	-
	0.97 - 1.00	-	-	-	-
	1.00 - 1.03	-10.776	-10.829	-11.033	-11.002
	$>$1.03	-0.891	-0.885	-0.894	-0.913
	All contracts$^{(a)}$	-1.365	-1.362	-1.380	-1.397
	DM	4.541**	5.309**	3.228**	n.a.

n.a. denotes not applicable.

- denotes no contracts.

(a) These average hedging errors are all significantly different from zero.

(b) Diebold-Mariano statistic computed for all contracts.

Table 14:

Hedging performance of alternative option price models constructed on the 18th of April 1995: Average over contracts of excess profits expressed in dollars relative to a risk-free investment. (** indicates significance at the 5% level.)

Day		GST-3	SNP	Mixture	Normal
18th	Investment (I)	402.606	403.003	397.517	396.991
	Moneyness (S/X)				
19th	<0.97	0.117	0.143	0.270	0.212
	0.97 - 1.00	0.085	0.110	0.238	0.252
	1.00 - 1.03	0.250	0.250	0.301	0.357
	>1.03	0.238	0.233	0.230	0.215
	All contracts[(a)]	0.228	0.226	0.246	0.247
	DM[(b)]	-4.602**	-5.310**	-0.441	n.a.
20th	<0.97	-	-	-	-
	0.97 - 1.00	-11.661	-11.625	-11.454	-11.519
	1.00 - 1.03	-4.001	-4.003	-3.982	-3.945
	>1.03	-9.776	-9.780	-9.795	-9.819
	All contracts[(a)]	-8.913	-8.916	-8.925	-8.940
	DM	4.617**	4.258**	4.530**	n.a.
21st	<0.97	-	-	-	-
	0.97 - 1.00	-	-	-	-
	1.00 - 1.03	-11.015	-11.081	-11.388	-11.210
	>1.03	-8.396	-8.400	-8.450	-8.468
	All contracts[(a)]	-8.424	-8.428	-8.481	-8.498
	DM	8.394**	8.844**	5.006**	n.a.

n.a. denotes not applicable.
- denotes no contracts.

(a) These average hedging errors are all significantly different from zero.

(b) Diebold-Mariano statistic computed for all contracts.

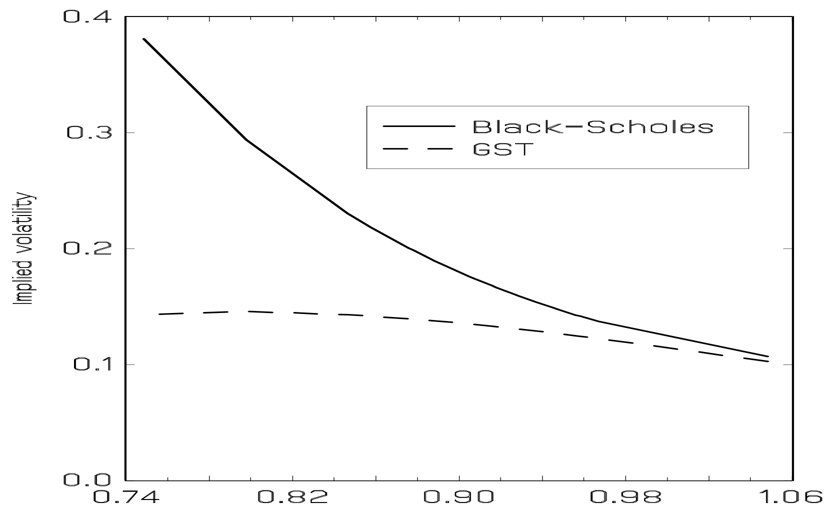

Figure 7: Volatility smiles for alternative models using data from April 4th, 1995, written on May options.

GST-3 model corrects the volatility skew arising from the Black-Scholes model, with implied volatility estimates now in a much narrower range of 10% to 14%. In fact, for the very deep in-the-money options, $0.75 < X/S < 0.9$, the implied volatility estimates are remarkably stable at around 13% to 14%, compared to the Black-Scholes implied volatilities which are range from 17% to 38%.

6 Conclusions

A general framework for pricing higher order moments and time-varying volatility in S&P500 options is developed. The approach consists of modelling the returns over the life of the option contract as a GST distribution. This yields a parametric form for the risk neutral density function which is used to price options. The parametric pricing model is shown to nest the Black-Scholes model.

The performance of a range of models is investigated using option contracts written on the S&P500 stock index for selected days in April 1995. The empirical results show significant gains to be made from pricing higher order moments in stock returns. The empirical significance of skewness, in particular, is consistent with the theoretical arguments provided by Bakshi, Kapadia and Madan (2003) for the importance of skewness in the pricing of options. The GST option price model corrects for volatility skews, thereby providing a consistent framework to price options in a single market across the full spectrum of moneyness. The GST modelling framework is also found to be superior, in general, to the alternative SNP and lognormal mixture models.

References

[1] Ait-Sahalia, Y. (1996), "Nonparametric Pricing of Interest Rate Derivative Securities", *Econometrica*, 64, 527-560.

[2] Ait-Sahalia, Y. (2002), "Telling From Discrete data Whether the Underlying Continuous-Time Model is a Diffusion", *Journal of Finance*, 57, 2075-2112.

[3] Ait-Sahalia, Y. and Lo, A.W. (1998), "Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices", *Journal of Finance,* 53, 499-547.

[4] Ait-Sahalia, Y. and Lo, A.W. (2000), "Nonparametric Risk Management and Implied Risk Aversion", *Journal of Econometrics,* 94, 9-51.

[5] Ait-Sahalia, Y. and Wang, Y. and Yared, F. (2001), "Do Option Markets Correctly Price the Probabilities of Movement of the Underlying Asset?", *Journal of Econometrics,* 102, 67-110.

[6] Ane, T. (1999), "Pricing and Hedging S&P500 Options with Hermite Polynomial Approximation: Empirical Tests of Madan and Milne's Model", *Journal of Futures Markets*, 19, 735-758.

[7] Bakshi, G., Cao, C. and Chen, Z. (1997), "Empirical Performance of Alternative Option Pricing Models", *Journal of Finance*, 52, 2003-2049.

[8] Bakshi, G., Kapadia, N. and Madan, D. (2003), "Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options", *The Review of Financial Studies,* 16, 101-143.

[9] Bates, D. (2000), "Post-87 Crash Fears in the S&P500 Futures Options", *Journal of Econometrics*, 94, 181-238.

[10] Bauwens, L. and Lubrano, M. (2002), "Bayesian Option Pricing Using Asymmetric GARCH Models", *Journal of Empirical Finance*, 9, 321-342.

[11] Black, F. and Scholes, M. (1973), "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, 81, 637-659.

[12] Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992), "ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence", *Journal of Econometrics*, 52, 5-59.

[13] Capelle-Blancard, G. Jurczenko, E. and Maillet, B. (2001), "The Approximate Option Pricing Model: Performances and Dynamic Properties", *Journal of Multinational Financial Management,* 11, 427-444.

[14] Chernov, M. and Ghysels, E. (2000), "A Study Towards A Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation", *Journal of Financial Economics*, 56, 407-458.

[15] Clement, E., Gourieroux, C. and Monfort, A. (2000), "Econometric Specification of the Risk Neutral Valuation Model", *Journal of Econometrics*, 94, 117-143.

[16] Corrado, C.J. and Su, T. (1997), "Implied Volatility Skews and Stock Index Skewness and Kurtosis implied by S&P 500 index Option Prices", *Journal of Derivatives*, Summer, 8-19.

[17] Diebold, F.X. and Mariano, R.S. (1995), "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, 13, 253-263.

[18] Duan, J-C. (1995), "The GARCH Option Pricing Model", *Mathematical Finance*, 5, 13-32.

[19] Duan, J-C. (1999), "Conditionally Fat-Tailed Distributions and the Volatility Smile in Options", *Draft Paper.*

[20] Dumas, B., Fleming, J. and Whaley, R. (1998), "Implied Volatility Functions: Empirical Tests", *Journal of Finance,* 53, 2059-2106.

[21] Dupire, B. (1994), "Pricing with a Smile", *Risk*, 7, 18-20.

[22] Engle, R.F. and Mustafa, C. (1992), "Implied ARCH Models From Option Prices", *Journal of Econometrics*, 52, 289-311.

[23] Fernandez, C. and Steele, M.F.J. (1998) "On Bayesian Modelling of Fat Tails and Skewness", *Journal of the American Statistical Association*, 93, 359-371.

[24] Forbes, C.S., Martin, G.M. and Wright, J. (2003), "Bayesian Estimation of a Stochastic Volatility Model Using Option and Spot Prices: Application of a Bivariate Kalman Filter", Draft Paper, Monash University.

[25] Gallant, R.A. and Tauchen, G. (1989), "Semi Non-Parametric Estimation of Conditionally Constrained Heterogeneous Processes: Asset Pricing Applications", *Econometrica,* 57, 1091-1120.

[26] Garcia, R. and Gencay, R. (2000), "Pricing and Hedging Derivative Securities with Neural Networks and a Homogeneity Hint", *Journal of Econometrics,* 94, 93-115.

[27] Hafner, C.M. and Herwartz, H. (2001), "Option Pricing Under Linear Autoregressive Dynamics, Heteroskedasticity and Conditional Leptokurtosis", *Journal of Empirical Finance*, 8, 1-34.

[28] Heston, S. (1993), "A Closed-Form Solution for Options with Stochastic Volatility with Application to Bond and Currency Options", *Review of Financial Studies*, 6, 327-343.

[29] Heston, S. and Nandi, S. (2000), "A Closed-Form GARCH Option Valuation Model", *Review of Financial Studies*, 13, 585-625.

[30] Hull, J.C. (2000), *Options, Futures, and Other Derivative Securities*, Fourth edition, Prentice Hall, New Jersey.

[31] Hull, J.C. and White, A. (1987), "The Pricing of Options on Assets with Stochastic Volatilities", *Journal of Finance*, 42, 281-300.

[32] Ingersoll, J.E.,Jr. (1987), *Theory of Financial Decision Making*, Rowman & Littlefield Studies in Financial Economics, USA.

[33] Jackwerth, J.C. (1999), "Option-Implied Risk-Neutral Distributions and Implied Binomial trees: A Literature Review", *The Journal of Derivatives*, Winter, 66-82.

[34] Jacquier, E. and Jarrow, R. (2000), "Bayesian Analysis of Contingent Claim Model Error", *Journal of Econometrics*, 94, 145-180.

[35] Jarrow, R. and Rudd, A. (1982), "Approximate Option Valuation for Arbitrary Stochastic Processes", *Journal of Financial Economics*, 10, 347-369.

[36] Kendall, M.G. and Stuart, A. (1969), *The Advanced Theory of Statistics*, Griffin, U.K.

[37] Lim, G.C., Lye, J.N., Martin, G.M. and Martin, V.L. (1998), "The Distribution of Exchange Rate Returns and the Pricing of Currency Options", *Journal of International Economics*, 45, 351-368.

[38] Lye, J.N. and Martin, V.L. (1993), "Robust Estimation, Non-normalities and Generalized Exponential Distributions", *Journal of the American Statistical Association*, 88, 253-259.

[39] Lye, J.N. and Martin, V.L. (1994), "Non-linear Time Series Modelling and Distributional Flexibility", *Journal of Time Series Analysis*, 15, 65-84.

[40] Lye, J.N., Martin, V.L. and Teo, L.E. (1998), "Parametric Distributional Flexibility and Conditional Variance Models with an Application to Hourly Exchange Rates", IMF Working Paper, 29.

[41] Martin, G.C., Forbes, C.S. and Martin, V.L. (2003), "Implicit Bayesian Inference Using Option Prices", Department of Econometrics and Business Statistics Working Paper 6/2003, Monash University.

[42] Melick, W.R. and Thomas, C.P. (1997), "Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil During the Gulf Crisis", *Journal of Financial Quantitative Analysis*, 32, 91-115.

[43] Pan, J. (2002), "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study", *Journal of Financial Economics*, 63, 3-50.

[44] Rosenberg, J.V. and Engle, R.F. (1997), "Option Hedging Using Empirical Pricing Kernels", NBER Working paper 6222, October.

[45] Rosenberg, J.V. (1998), "Pricing Multivariate Contingent Claims Using Estimated Risk-Neutral Density Functions", *Journal of International Money and Finance*, 17, 229-247.

[46] Sabbatini, M. and Linton, O. (1998), "A GARCH Model of the Implied Volatility of the Swiss Market Index from Option Prices", *International Journal of Forecasting*, 14, 199-213.