

ISSN 1440-771X
ISBN 0 7326 1074 5

Implicit Bayesian Inference Using Option Prices

Gael M Martin, Catherine S Forbes and Vance L Martin

Working Paper 5/2000

July 2000

**DEPARTMENT OF ECONOMETRICS
AND BUSINESS STATISTICS**



AUSTRALIA

Implicit Bayesian Inference Using Option Prices

Gael M. Martin, Catherine S. Forbes and Vance L. Martin

July, 2000

Abstract

A Bayesian approach to option pricing is presented, in which posterior inference about the underlying returns process is conducted implicitly, via observed option prices. A range of models which allow for conditional leptokurtosis, skewness and time-varying volatility in returns, are considered, with posterior parameter distributions and model probabilities backed out from the option prices. Fit, predictive and hedging densities associated with the different models are produced. Models are ranked according to several criteria, including their ability to fit observed option prices, predict future option prices and minimize hedging errors. In addition to model-specific results, averaged predictive and hedging densities are produced, the weights used in the averaging process being the posterior model probabilities. The method is applied to option price data on the S&P500 stock index. Whilst the results provide some support for the Black-Scholes model, no one model dominates according to all criteria considered.

Keywords: Bayesian Implicit Inference; Option Pricing Errors; Option Price Prediction; Hedging Errors; Nonnormal Returns Models; GARCH; Bayesian Model Averaging.

1 Introduction

This paper presents a Bayesian approach to option pricing, in which posterior inference is conducted implicitly, via observed option prices. The method is applied to option price data on the S&P500 stock index recorded on the last two trading days of 1995. A range of models for the underlying returns process are considered, with both posterior distributions for the parameters of those models and posterior model probabilities backed out from the option prices. The models considered allow for both time-varying conditional volatility, using the Generalized Autoregressive Conditional Heteroscedasticity (*GARCH*) framework introduced by Engle (1982) and Bollerslev (1986), and skewness and leptokurtosis in the conditional distribution of returns, using the framework of Fernandez and Steel (1998). The modelling framework nests the returns model underlying the option pricing model of Black and Scholes (1973), namely one of normality with constant volatility.

Fit and predictive densities associated with the different models are produced. These densities are used to assess the out-of sample performance of the different parametric models. The hedging performance of the different returns models is also assessed via the construction of posterior distributions for hedging errors. In addition to considering model-specific results, the concept of Bayesian model averaging is invoked, to produce averaged predictive and hedging distributions for the option prices. The weights used in the averaging process are the posterior probabilities associated with the different models. Results based upon implicit estimation are compared with results based upon direct estimation using returns data.

Although most of the statistical work on option prices is based on the classical paradigm, some Bayesian analyses have been performed. Boyle and Ananthanarayanan (1977) are amongst the first investigators to have applied Bayesian principles to option pricing, producing credible intervals for the Black-Scholes option price, based on a posterior distribution for the volatility parameter estimated from returns data. They allude to the possibility of the Bayesian approach being used to discriminate between different models for the returns process. Korolyi (1993) shows how Bayesian methods, again as based on returns data, can be used to improve the precision with which market option prices are estimated. Particular emphasis is given to the use of prior information on the cross-sectional patterns in return volatilities for groups of different types of stocks. Bayesian modelling of *GARCH* processes, with an application to the production of theoretical stock option prices is found in Bauwens and Lubrano (1998). More recently, Jacquier and Jarrow (2000) produce posterior distributions for the volatility parameter of the Black-Scholes model implicitly, via observed option prices. Deviations from the Black-Scholes assumptions are catered for non-parametrically.

Classical work which relates to the present paper includes Engle and Mustafa (1992), Bakshi, Cao and Chen (1997), Corrado and Su (1997), Lim et al (1998), Sabbatini and Linton (1998) and

Lim, Martin and Martin (2000). In Lim et al (1998), alternative models for exchange rate returns are formulated within the generalized Student t framework of Lye and Martin (1993). These models are estimated from the returns data using Maximum Likelihood, and the implications for option pricing of nonnormality and time-varying volatility investigated. In Corrado and Su (1997) and Lim, Martin and Martin (2000), option prices are used to estimate returns models which accommodate skewness and leptokurtosis. Significant option-implied negative skewness and excess kurtosis is found in both cases, with the Corrado and Su results pertaining to options on the S&P500 index and the Lim, Martin and Martin results pertaining to currency options. Bakshi, Cao and Chen (1997) conduct implicit estimation of alternative stochastic volatility models, via option prices on the S&P500 index. They also find evidence of departures from the Black-Scholes specifications and that the use of a more sophisticated option pricing model can reduce both pricing and hedging errors. Engle and Mustafa (1992) and Sabbatini and Linton (1998) estimate implied *GARCH* models from option prices, making comparisons with such models estimated directly from returns data.

An outline of the present paper is as follows. Section 2 discusses the application of the Bayesian statistical paradigm to option pricing. We follow Jacquier and Jarrow (2000) in specifying normally distributed option pricing errors. This distributional specification allows for the implicit estimation of posterior quantities related to the underlying returns models, from which fit, predictive and hedging distributions for the option prices themselves are to be constructed. In Section 3 the discussion is specialized, firstly to the Black-Scholes option pricing model. Parametric extensions to the Black-Scholes model are then introduced, with details given of the distributional framework used to model time-varying conditional volatility and nonnormality in the conditional distribution of returns. Section 4 outlines the criteria used to rank the different models, namely, implicit posterior probability, plus fit, predictive and hedging performance. In this section, we also outline the application of Bayesian model averaging, with the method of producing averaged predictive and hedging distributions described. In Section 5, the empirical results based on option price data on the S&P500 stock index are presented. Whilst the results provide some evidence favouring the Black-Scholes model, substantial option-implied excess kurtosis is estimated, with a Student- t model with constant volatility providing by far the best out-of-sample fit of observed option prices. The prediction and hedging performance of all models is remarkably similar, with no one model dominating. As a consequence, the model averaging process produces results which are also similar to those pertaining to the individual models. The hedging results suggest that all models considered are still misspecified up to a point. The comparison of the implicit and returns-based results suggests that option prices do contain information over and above that which is contained

in historical returns data and that this information is beneficial in reducing pricing errors. Some conclusions are drawn in Section 6.

2 Bayesian Inference in an Option Pricing Framework

The price of an option written on a non-dividend paying asset is the expected value of the discounted payoff of the option. For a European call option, the price is

$$q = E [e^{-r\tau} \max(S_T - K, 0)], \quad (1)$$

where the expectation is taken with respect to the risk neutral probability measure. The notation used in (1) is defined as follows:

- T = the time at which the option is to be exercised;
- τ = the length of the option contract;
- K = the exercise price;
- S_T = the spot price of the underlying asset at the time of maturity;
- r = the risk-free interest rate assumed to hold over the life of the option.

The option price is thus a function of certain observable quantities, namely r , K and τ . As the expectation is evaluated at time t , it is also a function of the observable level of the spot price prevailing at that time, S_t . Since the option price involves the evaluation of the expected payoff at the time of maturity, the price depends on (i) the assumed stochastic process for S , or alternatively, on the assumed distribution for returns on the asset; and (ii) the values assigned to the unknown parameters of that underlying process. In this paper, we explicitly allow for the uncertainty associated with both (i) and (ii), by producing posterior probability distributions for the model-specific parameters and posterior probabilities for a range of alternative models. We integrate out parameter uncertainty by producing predictive densities for out-of-sample option prices, for each of the different models. We then integrate out model uncertainty by producing model-averaged predictive and hedging densities.

Posterior inferences are to be produced implicitly from observed market option prices¹. For this to occur, option prices need to be assigned a particular distributional model. In the spirit of Jacquier and Jarrow (2000), we consider the following representation of an option pricing model with unobservable pricing error u_i :

$$C_i = \beta q(z_i, \theta) + u_i, \quad (2)$$

¹As noted in the introduction, in the empirical section we also produce inferences directly from returns data. However, the focus of the paper is on implicit estimation.

where C_i is the observed market price of, say, a call option on an underlying financial asset and $z_i = (r, K, \tau, S)'$ is the vector of observable factors which determine the option price. The index i indicates variation over time, as well as cross sectional variation across different option contracts at a given point in time. That is, (2) is a model appropriate for a panel of option price data. The function $q(\cdot, \cdot)$ in (2) represents the functional form of the theoretical option price which, in turn, reflects the distribution assumed for the underlying returns process. Since any option price formula involves the evaluation of an expectation with respect to the (risk-neutral) distribution of the underlying asset, $q(\cdot, \cdot)$ is a function of the parameters which characterize that distribution, θ .

The error term in (2), u_i , is the error between the observed market option price and the theoretical option price, with u_i assumed to have a non-degenerate distribution. In this paper, we assume that $u_i \sim N(0, \sigma_u^2)$ for all i , truncated from below, since the option price must satisfy a particular no-arbitrage lower bound. The use of alternative distributional assumptions for u_i is left for future work². The presence of u_i reflects the reality that any option model, $q(\cdot, \cdot)$, no matter how complex, is only an approximation of the process which has lead to the determination of an observed option price. That is, u_i encompasses ‘model error’. It may also reflect ‘market error’, in which an observed option price differs from its theoretical counterpart as the result of the non-synchronous recording of spot and option prices, transaction costs and other market frictions. As noted by Jacquier and Jarrow (2000), these two components of u_i are not identifiable without further assumptions and no attempt is made in the present paper to identify them separately.

Given the assumption of normality for u_i , the likelihood function for $\Theta = (\theta, \beta, \sigma_u)$ is given by:

$$L(\Theta) \propto \sigma_u^{-N} \exp\left(-\frac{1}{2\sigma_u^2} \sum_{i=1}^N [C_i - \beta q(z_i, \theta)]^2\right), \quad (3)$$

where N is the number of option prices in the data set. Given a prior density on Θ , $p(\Theta)$ ³, and defining the $(N \times 1)$ vector of observed option prices as \mathbf{c} , the joint posterior density for Θ is given by:

$$p(\Theta|\mathbf{c}) \propto L(\Theta) \times p(\Theta). \quad (4)$$

Given the nonstandard nature of (4), which obtains even for the simplest case of the Black-Scholes model, numerical procedures are required in order to produce both marginal posterior distributions for the elements of Θ and the marginal likelihood associated with the specified model, to be used

²Jacquier and Jarrow (2000) also confine their attention to a normally distributed u_i . However they generalize the variance structure by allowing for limited amount of heteroscedasticity in u_i , to reflect the large variation in the size of the option prices across the moneyness spectrum. They also consider option pricing models which are linear in logs as well as linear in levels. Since our analysis is complicated by the introduction of alternative returns models, we have chosen to focus only on the simplest of specifications for u_i .

³The prior density can be specified in a way which reflects information from the underlying returns data. We return to this point in Section 5.

in the construction of relative model probabilities and in the model averaging. As well as enabling posterior quantities relating to the underlying returns model to be produced, the numerical procedures enable fit, predictive and hedging distributions to be estimated, as will be detailed in Section 4.

3 Alternative Option Pricing Models

The adoption of alternative distributional assumptions about returns means that the functional form for the theoretical option price, $q(z_i, \theta)$, in (2) alters, as does the specification of the parameter vector θ . In this section, we outline the alternative distributions adopted for returns and describe their implications for $q(z_i, \theta)$. We begin by looking at the Black-Scholes option pricing model, based on the assumption of a normal returns process with constant volatility. We then look at extended models which allow for both nonnormalities and time-varying conditional volatility.

3.1 The Black-Scholes Model

The assumption that the risk-neutral stochastic process for the spot price is

$$d \ln S = (r - 0.5\sigma^2) dt + \sigma dW, \quad (5)$$

where

$$dW \sim N(0, dt), \quad (6)$$

yields the Black-Scholes option price,

$$q(z_i, \sigma) = S\Phi(d_1) - Ke^{-r\tau}\Phi(d_2), \quad (7)$$

where

$$d_1 = \frac{\ln(S/K) + (r + \sigma^2/2)\tau}{\sigma\sqrt{\tau}},$$

$$d_2 = \frac{\ln(S/K) + (r - \sigma^2/2)\tau}{\sigma\sqrt{\tau}}$$

and $\Phi(x)$ denotes the cumulative normal distribution function evaluated at x ; see, for example, Hull (2000). The parameter σ represents the instantaneous volatility of returns on the asset. The Black-Scholes option price for a dividend paying asset, like the S&P500 index for example, is given by (7) with S replaced by $S - D$, where D represents the present value of dividends to be paid over the life of the option.

The expression in (7) makes clear the analytical relationship which exists between the single unknown parameter, $\theta = \sigma$, and the option price $q(z_i, \sigma)$, in the Black-Scholes framework. It is

via this analytical expression that the volatility parameter, σ , enters the joint posterior density for $\Theta = (\sigma, \beta, \sigma_u)$ in (4). Numerical procedures are used to produce simulated values from $p(\Theta|\mathbf{c})$, as well as an estimate of the implied marginal posterior for volatility, $p(\sigma|\mathbf{c})$. This density can be compared both with the volatility densities associated with more highly parameterized returns models, as well as being contrasted with the volatility distributions produced directly from returns data.

3.2 Non-Black-Scholes Models

3.2.1 Motivations for the Generalization of the Black-Scholes Model

Two key assumptions of the Black-Scholes option price specified in (7) are that returns are normally distributed and that the volatility of returns, σ , is constant over the life of the contract. However, there is much empirical evidence that both of these assumptions are invalid. Friedman and Vandersteel (1982) and Lim et al (1998), amongst others, find that the distribution of exchange rate returns exhibits excess kurtosis relative to the normal distribution. Baillie and Bollerslev (1989, 1991), Harvey, Ruiz and Sentana (1992), Harvey, Ruiz and Shephard (1994) and Kim, Shephard and Chib (1998) find evidence of significant time varying volatility; see also Bollerslev, Chou and Kroner (1992) for a review of this literature. De Vries (1994), Lim et al (1998) and Lye, Martin and Teo (1998) find evidence of skewness in returns data.

As has also been widely documented, Black-Scholes implied volatilities are not constant across strike prices or maturity (see, for example, Rubenstein, 1994, Bakshi, Cao and Chen, 1997, Corrado and Su, 1997 and Dumas, Fleming and Whaley, 1998). Specifically, implied volatility ‘smiles’ or ‘smirks’ across strike prices, which, in turn, vary in intensity depending on the time to expiration, have become a stylized fact in empirical work on option prices. Such patterns have been shown to be evidence of implied returns models which deviate from the specifications of the Black-Scholes model; see, for example, Hull and White (1987), Johnson and Shanno (1987), Scott (1987), Wiggins (1987), Corrado and Su (1997) and Lim, Martin and Martin (2000).

In this paper, we extend the distributional assumptions of the Black-Scholes model using the distributional framework of Fernandez and Steel (1998), augmented by a time varying conditional variance structure. In the next sub-section, we describe this modelling framework and discuss briefly how it can be used in implicit estimation using option prices. Details of the numerical procedures applied in the implementation of the approach are provided in Section 5.

3.2.2 Conditional Leptokurtosis, Skewness and Time-Varying Volatility in Returns

Defining S_t as the spot price of the underlying asset at the end of day t , we initially assume that continually compounded daily returns on the asset, $y_t = \ln(S_t/S_{t-1})$, can be modelled as

$$y_t = \mu + \sigma e_t, \quad (8)$$

where e_t is a standardized error term given by

$$e_t \sim (0, 1). \quad (9)$$

The parameter σ denotes the standard deviation of daily returns. Assuming 250 trading days in the year, an annualized measure of the volatility of returns is given by $\sqrt{250}\sigma$. The parameter μ denotes the mean return. With the risk-neutral distribution being the distribution with respect to which the option price in (1) is defined, $\mu = r - 0.5\sigma^2$ is the mean formulation which enters the specification of the option price; see Hull (2000).

Four alternative distributional forms are assumed for the error term e_t in (9): normal, Student t , skewed normal and skewed Student t . The two skewed distributions are specified using the framework of Fernandez and Steel (1998). This framework nests the symmetric normal and Student t distributions. Denoting by z_t a random variable with mean μ_z and variance σ_z^2 , Fernandez and Steel allow for skewness in the distribution of z_t by defining a density function, $p_f(z_t|\gamma)$, as

$$p_f(z_t|\gamma) = \frac{2}{\gamma + \frac{1}{\gamma}} \left\{ f\left(\frac{z_t}{\gamma}\right) I_{[0,\infty)}(z_t) + f(\gamma z_t) I_{(-\infty,0)}(z_t) \right\}, \quad (10)$$

where $f(\cdot)$ is a symmetric density function with a single mode at zero and $I_A(z_t)$ denotes the indicator function for the set A . The mean and variance of z_t are defined respectively as

$$\mu_z = \left(\frac{\gamma^2 - 1/\gamma^2}{\gamma + 1/\gamma} \right) \int_0^\infty x 2f(x) dx$$

and

$$\sigma_z^2 = \left(\frac{\gamma^3 + 1/\gamma^3}{\gamma + 1/\gamma} \right) \int_0^\infty x^2 2f(x) dx - \mu_z^2.$$

The parameter γ denotes the degree of skewness in the distribution, with $\gamma > 1$ corresponding to positive skewness, $\gamma < 1$ corresponding to negative skewness and $\gamma = 1$ corresponding to symmetry. The density $p_f(z_t|\gamma)$ retains the mode of zero, with the sign and magnitude of $\gamma - 1$ determining the sign and magnitude of μ_z . When $f(\cdot)$ defines the normal density function, z_t is referred to as having a skewed normal distribution, $SN(\mu_z, \sigma_z^2, \gamma)$. When $f(\cdot)$ defines the Student t density function, z_t is referred to as having a skewed Student t distribution, $SSt(\mu_z, \sigma_z^2, \gamma, \nu)$. When $\gamma = 1$,

these skewed distributions collapse respectively to the standard normal distribution, $N(0, 1)$, and the Student t distribution, $St(0, \frac{\nu}{\nu-2}, \nu)$, respectively.

Using the distributional framework in (10), we can define the alternative specifications for returns, y_t , in (8) as corresponding to alternative specifications for $f(\cdot)$ and γ , as follows:

$$\begin{aligned}
M_1 : f(\cdot) \text{ normal}; \quad \gamma = 1; & \quad \left(\frac{y_t - \mu}{\sigma} \right) \sim N(0, 1) \\
M_2 : f(\cdot) \text{ Student } t; \quad \gamma = 1; & \quad \sqrt{\frac{\nu}{\nu-2}} \left(\frac{y_t - \mu}{\sigma} \right) \sim St(0, \frac{\nu}{\nu-2}) \\
M_3 : f(\cdot) \text{ normal}; \quad \gamma \neq 1; & \quad \mu_z + \sigma_z \left(\frac{y_t - \mu}{\sigma} \right) \sim SN(\mu_z, \sigma_z^2, \gamma) \\
M_4 : f(\cdot) \text{ Student } t; \quad \gamma \neq 1; & \quad \mu_z + \sigma_z \left(\frac{y_t - \mu}{\sigma} \right) \sim SSt(\mu_z, \sigma_z^2, \gamma, \nu)
\end{aligned} \tag{11}$$

Model M_1 specifies returns as being normally distributed with constant volatility, σ . It thus corresponds to the discrete time version of the returns model which underlies the Black-Scholes option price. We shall subsequently refer to M_1 as the Black-Scholes model. Model M_2 specifies standardized returns as being $St(0, 1, \nu)$, thereby accommodating the leptokurtosis which is a feature of returns data. The degrees of freedom parameter, ν , controls the degree of kurtosis.⁴ Model M_3 allows for skewness in returns, whilst model M_4 allows for both leptokurtosis and skewness.

Models M_2 to M_4 , whilst extending the distributional assumptions of the Black-Scholes model, maintain its specification of constant volatility. With volatility clustering and an associated correlation structure in squared returns, being a further well-documented feature of financial data, it is appropriate to generalize the returns model in (8) to allow for time-varying conditional variance. As such, (8) is rewritten as

$$y_t = \mu + \sigma_t e_t, \tag{12}$$

where σ_t is specified to have a first order Generalized Autoregressive Conditional Heteroscedastic (*GARCH*(1, 1)) structure; see Bollerslev (1986),

$$\sigma_t^2 = a + b(y_{t-1} - \mu)^2 + c\sigma_{t-1}^2. \tag{13}$$

With the distribution of

$$e_t = \frac{y_t - \mu}{\sigma_t}, \tag{14}$$

conditional on past returns, allowed to assume the four forms listed in (11), the generalization of the returns process to (12) produces four additional models, which we denote as follows:

⁴Note that the specification of $e_t = (y_t - \mu)/\sigma$ as a *standardized* Student t variate enables the peakedness as well as the fat-tailed features of returns data to be captured.

$$\begin{aligned}
M_5 : f(.) \text{ normal}; \quad \gamma = 1; \quad \sigma_t \text{ GARCH} & \quad \left(\frac{y_t - \mu}{\sigma_t} \right) \sim N(0, 1) \\
M_6 : f(.) \text{ Student } t; \quad \gamma = 1; \quad \sigma_t \text{ GARCH} & \quad \sqrt{\frac{\nu}{\nu-2}} \left(\frac{y_t - \mu}{\sigma_t} \right) \sim St(0, \frac{\nu}{\nu-2}, \nu) \\
M_7 : f(.) \text{ normal}; \quad \gamma \neq 1; \quad \sigma_t \text{ GARCH} & \quad \mu_z + \sigma_z \left(\frac{y_t - \mu}{\sigma_t} \right) \sim SN(\mu_z, \sigma_z^2, \gamma) \\
M_8 : f(.) \text{ Student } t; \quad \gamma \neq 1; \quad \sigma_t \text{ GARCH} & \quad \mu_z + \sigma_z \left(\frac{y_t - \mu}{\sigma_t} \right) \sim SSt(\mu_z, \sigma_z^2, \gamma, \nu)
\end{aligned} \tag{15}$$

Models M_1 to M_8 all imply a different functional form for the theoretical option price, $q(z_i, \theta)$, in (2), as well as a different specification for the parameter vector, θ .⁵ For all models other than M_1 , $q(.,.)$ does not have a closed-form solution. Rather, the expectation implicit in $q(.,.)$ must be computed using numerical methods. The approach adopted in this paper is to simulate S over the life of the contract according to a discretized version of the stochastic process in (5); see Gouriéroux and Monfort (1994), for a general discussion of simulating continuous time SDEs. Depending on the assumed model, the random increments in (5) are generated from the relevant standardized distribution: $N(0, 1)$ for M_5 , $St(0, 1, \nu)$ for M_2 and M_6 ; $SN(0, 1, \gamma)$ for M_3 and M_7 ; and $SSt(0, 1, \gamma, \nu)$ for M_4 and M_8 . For M_5 to M_8 , a $GARCH(1, 1)$ process is simulated over the life of the option, rather than constant volatility over the life of the option being imposed. For each model M_2 to M_8 , simulation of the relevant process for S is repeated h times, producing $S_T^{(l)}$, $l = 1, 2, \dots, h$, and the expectation in (1) approximated by the sample mean of $e^{-r\tau} \max(S_T^{(l)} - K, 0)$. In the case of a dividend-earning asset, $S_T^{(l)}$ is replaced by $S_T^{(l)} - D$, as explained earlier. It is via this estimated option price that the parameters of the extended models enter the joint posterior density for Θ in (4). Both antithetic and control variates are used to reduce the simulation error; see, for example, Hull (2000). The analytical Black-Scholes option price, $q(z_i, \sigma)$, is used as the control variate.⁶

4 Model Ranking and Averaging

The various models considered for returns are to be ranked according to several different criteria. In this section, these criteria are discussed, prior to their application in the empirical section. As well as ranking the individual models, model-averaged predictive and hedging results are to be produced. The mechanics of producing these averaged results are also discussed in this section.

⁵In the case of M_1 , y_t is normal, thereby yielding a lognormal distribution for the terminal index value S_T , conditional on S_t . Imposing the restriction that $\mu = r - 0.5\sigma^2$ in (8), yields, as already noted in the text, the risk neutral probability distribution, which is the appropriate distribution for evaluating the expectation in (1). For the remaining set of models, M_2 to M_8 , the associated distribution of S_T can be shown to be the pertinent risk neutral distribution for pricing options provided that the mean of the returns distribution is specified as $\mu = r - 0.5\sigma^2$ for M_1 to M_4 , and as $\mu = r - 0.5\sigma_t^2$ for M_5 to M_8 ; see Jarrow and Rudd (1982) and Bollerslev and Mikkelsen (1999).

⁶In Lim, Martin and Martin (2000), an alternative method, based on numerical integration of the option payoff with respect to alternative distributions for S_T , is presented.

4.1 Implicit Bayes Factors and Model Probabilities

Given K alternative models considered for returns, M_k , $k = 1, 2, \dots, K$, K posterior probabilities, $P(M_k|\mathbf{c})$, are to be computed from the option price data, with

$$\sum_{k=1}^K P(M_k|\mathbf{c}) = 1. \quad (16)$$

The model probabilities are used to rank the K models. These probabilities are derived from the posterior odds ratios, constructed for each model, M_2, M_3, \dots, M_K , relative to a reference model, M_1 . Denoting the vector of parameters specific to model M_k by $\Theta_k = (\theta_k, \beta, \sigma_u)$, the posterior odds ratio for M_k versus M_1 is given by

$$\begin{aligned} \frac{P(M_k|\mathbf{c})}{P(M_1|\mathbf{c})} &= \frac{P(M_k)}{P(M_1)} \times \frac{p(\mathbf{c}|M_k)}{p(\mathbf{c}|M_1)} \\ &= \text{Prior Odds} \times \text{Bayes Factor}, \end{aligned} \quad (17)$$

where

$$p(\mathbf{c}|M_k) = \int_{\Theta_k} L(\Theta_k|M_k)p(\Theta_k|M_k)d\Theta_k \quad (18)$$

is the marginal likelihood of M_k , with $L(\Theta_k|M_k)$ and $p(\Theta_k|M_k)$ respectively denoting the likelihood and prior under M_k .⁷ The model probabilities are calculated by solving the $(k - 1)$ ratios in (17) subject to (16) and the models then ranked as *a posteriori* most probable to least probable according to the size of the probabilities.

Given the complexity of the problem, closed form solutions for the marginal likelihoods are not available. We choose to estimate the marginal likelihoods and, hence, the Bayes Factors, using the method of Chib (1995). Consider this method as applied to estimating the marginal likelihood for M_k in (18). For any point in the space of Θ_k , $p(\mathbf{c}|M_k)$ can be expressed as

$$p(\mathbf{c}|M_k) = \frac{L(\Theta_k|M_k)p(\Theta_k|M_k)}{p(\Theta_k|\mathbf{c},M_k)}. \quad (19)$$

Whilst the likelihood and prior can be evaluated at any point in Θ_k , the posterior density, $p(\Theta_k|\mathbf{c},M_k)$, cannot, due to its nonstandard form. However, an estimate of the posterior ordinate can be readily obtained using the simulation output.

4.2 Fit Performance

For any model M_k with parameter vector θ_k , the distribution of the residual associated with fitting the *ith* option price, C_i ,

$$res_i^{(j)} = C_i - \beta^{(j)}q(z_i, \theta_k^{(j)}), \quad (20)$$

⁷Note that parameters β and σ_u are common to each Θ_k , $k = 1, 2, \dots, K$.

is generated via the simulated values of $\beta^{(j)}$ and $\theta_k^{(j)}$. A density function for res_i could then be estimated using some form of kernel smoothing. This approach, however, is computationally expensive, as it needs to be performed for all i and for all models. As an alternative, the mean, \overline{res}_i , and standard deviation, $sd(res_i)$, of a given set of res_i values is used to construct the interval,

$$INT_i = \overline{res}_i \pm 1.96 \times sd(res_i). \quad (21)$$

In the case where the empirical distribution of (20) is well approximated by a normal distribution, the interval in (21) represents the 95% Highest Posterior Density (HPD) interval for res_i .

Given a sample of N option prices and, hence, N values of INT_i for model M_k , the proportion of INT_i intervals which cover zero can be calculated. The best fitting model so defined is the model for which this proportion is the highest. If the observed option prices used in (20) are the same prices from which the posterior density function for Θ_k is estimated, the fit assessment is *within-sample*. If not, the fit assessment is *out-of-sample*. The latter is the form of fit assessment used in the empirical section.

4.3 Predictive Performance

For any model M_k with parameter vector Θ_k , the predictive density for an out-of-sample option price, C_f say, is given by:

$$p(C_f|\mathbf{c}) = \int p(C_f|\mathbf{c}, \Theta_k) p(\Theta_k|\mathbf{c}) d\Theta_k, \quad (22)$$

where $p(C_f|\mathbf{c}, \Theta_k)$ is a truncated normal density, given the assumption of a truncated normal distribution for u_f in (2). Repeated draws from $p(\Theta_k|\mathbf{c})$, $\Theta_k^{(j)}$, $j = 1, 2, \dots, B$, can be used to construct repeated conditional densities, $p(C_f|\mathbf{c}, \Theta_k^{(j)})$, which can then be averaged to produce an estimate of $p(C_f|\mathbf{c})$ as

$$p(\widehat{C}_f|\mathbf{c}) = \frac{1}{B} \sum_{j=1}^B p(C_f|\mathbf{c}, \Theta_k^{(j)}). \quad (23)$$

For the non Black-Scholes models, the function through which $p(C_f|\mathbf{c}, \Theta_k^{(j)})$ depends on $\theta_k^{(j)}$, namely $q(z_f, \theta_k^{(j)})$, is itself estimated using the simulation procedure described in Section 3.2.2.⁸ If C_f is one of the elements of \mathbf{c} , then (23) provides a method of assessing the *within-sample* predictive performance of M_k . Otherwise, the assessment is one of *out-of-sample* predictive performance. As with fit performance, the latter is the form of assessment used in the empirical section.

The predictive density in (23) is used to rank the predictive performance of the models in several different ways. First, (23) is used to assign a probability to the observed bid-ask spread associated

⁸Note that the theoretical option price associated with C_f must be expressed as a function of the same option contract specifications, as given in z_f , as does C_f itself.

with option contract for which C_f is the market price.⁹ This calculation is repeated for all option contracts included in the hold-out sample, the predictive probability recorded for model M_k being the average across all probabilities computed. Secondly, with the predictive mode taken as a point predictor of C_f , the accuracy of each model is assessed in terms of the proportion of predictive modes which fall within the observed bid-ask spreads. The same calculation is performed for the predictive means. Thirdly, the proportion of bid-ask spreads which fall within the interquartile range and the 95% probability interval associated with (23) is calculated for each model. Finally, the proportion of observed market prices (= the average of the bid and ask prices) which fall within these same ranges is also calculated.¹⁰

4.4 Hedging Performance

Another measure of the performance of alternative option price models is the extent to which the associated hedging errors deviate from zero. In this paper attention is restricted to delta hedges; see Hull (2000) for a discussion of alternative hedging schemes. At time t , the delta for the i th option price, based on the assumption that M_k describes the returns process, is defined as

$$\delta_t^k = \frac{\partial q(z_i, \theta_k)}{\partial S_t}. \quad (24)$$

In computing the hedging errors, the portfolio at time t consists of going short in the option and long in the underlying asset by an amount of δ_t^k shares in the asset, and investing the residual, $C_t - \delta_t^k S_t$, at the risk free interest rate r . At time $t + \Delta t$, the hedge error associated with model M_k , based on the portfolio at t , is given by; see Bakshi, Cao and Chen (1997)

$$H_k = \delta_t^k (S_{t+\Delta t} - S_t e^{r\Delta t}) - (C_{t+\Delta t} - C_t e^{r\Delta t}). \quad (25)$$

The posterior distribution of the hedging error in (25) is derived from the posterior distribution for the parameters of model M_k , via δ_t^k . In fact, the distribution of H_k is a simple translation of the distribution of δ_t^k , obtained by recentering this distribution by $C_{t+\Delta t} - C_t e^{r\Delta t}$, and rescaling it by $S_{t+\Delta t} - S_t e^{r\Delta t}$. Thus, the hedge distribution can be generated by evaluating δ_t^k at the simulated values of $\theta_k^{(j)}$, and using a nonparametric kernel to estimate the density.

⁹With regard to the S&P500 option quote price data to be analysed in Section 5, there is usually only one bid-ask spread associated with a particular option contract (where the specification of that contract includes the current spot price of the index). For some contracts, however, there are several bid-ask spreads quoted. These spreads are averaged over before used in the probability calculation described in the text.

¹⁰Note that there is a large literature on the market-related factors which influence the bid-ask spreads associated with option prices. In particular, attempts have been made to explain the way in which the spreads vary across different type of option contracts; see, for example George and Longstaff (1993). On the assumption that these factors do not relate to the nature of the underlying returns process, it is legitimate to ignore them and simply take the observed spreads as given intervals to which the different models assign varying predictive probabilities. This assumption may be questionable however; see, for instance, Cho and Engle (1999).

In the case where the price of the underlying asset, S , follows the generating process given by (5) and (6), and the Black-Scholes model is used to specify δ_t , the hedging distribution becomes degenerate at zero as $\Delta t \rightarrow 0$. This is the case where the portfolio is rebalanced instantaneously. In practice, rebalancing is performed at discrete points in time thereby resulting in a non-degenerate hedging distribution even when the Black-Scholes option pricing model is correct. However, hedging errors closely distributed around zero would represent evidence of the validity of the Black-Scholes model.

In the case where S follows a process which is consistent with any one of the non-Black-Scholes returns models considered in the paper, M_2 to M_8 , and this same model is used to specify δ_t , hedging errors closely distributed around zero can be viewed as evidence in favour of this model. The alternative models can thus be ranked according to the nature of their hedging distributions.¹¹

4.5 Model Averaging

As described in Section 4.1, the posterior probability of each model M_k , $P(M_k|\mathbf{c})$, $k = 1, 2, \dots, K$, can be backed out of the option prices. These posterior probabilities can be used to produce averaged predictive and hedging densities as follows.¹² Consider the case of predictive averaging first. Given model specific predictive densities, $p(C_f|M_k, \mathbf{c})$, $k = 1, 2, \dots, K$, the averaged predictive density, $p_a(C_f|\mathbf{c})$, is defined as:

$$p_a(C_f|\mathbf{c}) = \sum_{k=1}^K p(C_f|M_k, \mathbf{c})P(M_k|\mathbf{c}). \quad (26)$$

This predictive density is then used to produce the various coverage statistics described earlier. If this averaging process is a valid representation of the behaviour of the option market participants, all of whom may price options according to different assumed models, the coverage statistics produced from (26) may well be larger than those produced from the predictive density associated with any one particular model.

The averaged hedging density is defined as:

$$p_a(H|\mathbf{c}) = \sum_{k=1}^K p(H_k|M_k, \mathbf{c})P(M_k|\mathbf{c}). \quad (27)$$

Note that H_k for each model M_k is a given function of the model specific θ_k , via (25). In this sense, the averaging process in (27) is different from that in (26), in which the argument of the densities being averaged, C_f , is the same for all models.

¹¹In computing the delta for the Black-Scholes model, an analytical solution is given by $\delta_t^{BS} = \Phi(d_1)$, where d_1 is defined in (7). For the other models, the derivative in (24) is computed numerically. To improve the accuracy of the numerical differentiation, a control variate is used for the non-Black-Scholes models based on the difference between the Black-Scholes analytical and numerical derivatives.

¹²See Raftery, Madigan and Hoeting (1997) and Geweke (1999) for discussions of the principles of Bayesian model averaging.

Table 1:
S&P500 Option Price Dataset

Moneyiness (S_t/K)	Average Market Price	Average Bid-Ask Spread	No. of Prices
Panel A: Estimation Dataset: 28/12/1995			
<i>OTM</i> : < 0.97	\$2.75	\$0.33	115
<i>ATM</i> : $0.97 - 1.03$	\$16.40	\$0.74	237
<i>ITM</i> : ≥ 1.03	\$81.00	\$1.00	752
			Total 1104
Panel B: Out of Sample Dataset: 29/12/1995			
<i>OTM</i> : < 0.97	\$2.84	\$0.35	45
<i>ATM</i> : $0.97 - 1.03$	\$16.09	\$0.77	113
<i>ITM</i> : ≥ 1.03	\$56.13	\$1.00	255
			Total 413

5 Empirical Application to S&P500 Option Prices

5.1 Data Description

The data used in the implicit estimation of the returns models is a sample of intraday prices on call options written on the S&P500 stock index.¹³ The prices relate to options traded on the 28th of December, 1995, with an expiry date of the 15th of March, 1996. A total of 1104 option prices are included in the sample, for option contracts extending over the full moneyness spectrum. Defining $S_t - K$ as the time- t intrinsic value of a call option, we follow Bakshi, Cao and Chen (1997) by categorizing options for which $S_t/K \in (0.97, 1.03)$ as *at-of-the-money (ATM)*, those for which $S_t/K \leq 0.97$ as *out-of-the-money (OTM)*, and those for which $S_t/K \geq 1.03$ as *in-the-money (ITM)*. Given a maturity length of approximately two and a half months, the options in the sample can be classified as medium-term. Each record in the data set comprises a bid-ask quote, the synchronously recorded spot price of the index, the time at which the quote was recorded, and the

¹³The data has been obtained from the Berkeley Options Database.

strike price. Since dividends are paid on the S&P500 index, in the option price formulae, the current spot price, S_t , is replaced by the dividend-exclusive spot price, $S_t - D$, where D is the present value of dividend paid over the life of the option.¹⁴ Only observations for which the average of the bid and ask prices exceeds the lower bound of $\max\{0, S_t - D - e^{-rt}K\}$ 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 nonsynchronicity between the spot and option prices.¹⁵ Panel A in Table 1 summarizes the main characteristics of the data set.

In assessing the fit, predictive and hedging performance of the different models, option price data for the 29th of December, 1995 is used. This data relates to option contracts with the same maturity date as described for the estimation data set and with the same exclusions having been applied. A total of 413 option prices are used to assess the out-of-sample performance of the models. The characteristics of this data set are summarized in Table 1, Panel B. As is evident, the characteristics of both datasets are very similar, apart from the average price of the *ITM* options, which is lower for the out-of sample dataset.¹⁶

5.2 Implicit Posterior Distributions for the Parameters of the Different Returns Models

5.2.1 Estimation Details

The first step in the implicit analysis is to produce estimates of the posterior distributions for the parameters of the alternative models. Of the eight different models described in Section 3.2.2, results are produced for only six: M_1 to M_6 . The skewed Normal-*GARCH* (M_7) and skewed Student *t-GARCH* (M_8) models are not estimated, since little evidence in favour of skewness is found in the constant volatility models. Also, since the focus is on the different returns models as implied by the option prices, we simplify the analysis by imposing $\beta = 1$ and $\sigma_u = 160$ in (2), for each returns model. The estimate of σ_u is obtained via a preliminary application of Maximum Likelihood to the likelihood function in (3), with $\beta = 1$ and $\theta = \sigma$. This simplification means that $\Theta_k = \theta_k$ for all $k = 1, 2, \dots, 6$, and that the joint posterior for model k can be expressed as

$$p(\theta_k | \mathbf{c}) \propto L(\theta_k) \times p(\theta_k), \quad (28)$$

with $L(\theta_k)$ as defined in (3) with $\beta = 1$ and $\sigma_u = 160$ imposed.

¹⁴Daily dividend data for 1995 and 1996 was obtained from Standard and Poors.

¹⁵Although the synchronous recording of spot and option prices means that the spot prices recorded in the sample vary over the day, we continue to use only a single subscript t to denote a spot price recorded on day t .

¹⁶The imposition of the lower bound had the effect of deleting from the original data set about 100 of the highest *ITM* option prices, thereby causing a decline in the average *ITM* price.

In the case of the Black-Scholes model (M_1), $q(z_i, \theta_1)$ is an analytical function of a single parameter $\theta_1 = \sigma$. As such, the posterior density in (28) can be evaluated over a grid of values for σ and normalized via one-dimensional numerical integration. A sample of independent σ values can be produced via the inverse cumulative distribution function technique. For all models other than M_1 , $q(z_i, \theta_k)$ is estimated via simulation. The need to estimate the option price, for each value of θ_k in the parameter space of the non-Black-Scholes models, is computationally expensive. If this computation were to be embedded within a Markov chain Monte Carlo scheme, the computational cost of the methodology would be prohibitive. As a consequence, the method is implemented in a way which relies on only bivariate independent sampling. For example, in the case of M_2 , the bivariate posterior for $\theta_2 = \{\sigma, \nu\}$ in (28) is evaluated over a grid of values for θ_2 and normalized using two-dimensional numerical integration. The marginal posteriors are, in turn, produced via a further application of one-dimensional numerical integration. The estimated marginals are used to produce the summary measures for each parameter. Once the joint posterior is evaluated, the cumulative distribution function is estimated and used to produce an independent sample of θ_2 values via the inverse of the estimated distribution function.

The estimation method described above for M_2 is directly applicable to M_3 , with $\theta_3 = \{\sigma, \gamma\}$. In the case of the larger models, the bivariate approach is maintained by fixing certain parameters at values estimated via preliminary analyses. For example, in the case of M_4 , γ is fixed at the marginal mean for γ estimated in M_3 . In the case of the Normal-*GARCH* model, M_5 , a limited Markov Chain Monte Carlo scheme, based upon iterative sampling from the conditional posteriors of a and $\{b, c\}$, is used to produce an initial estimate of a . This value is then fixed, and an independent sample of $\{b, c\}$ produced from the joint posterior for $\{b, c\}$, conditional on the fixed value of a . In models M_5 and M_6 , a is held fixed at this same value. In M_6 , ν is specified at the value estimated for M_2 .

5.2.2 Priors

Estimates of the marginal posteriors are produced using informative priors for ν and γ and non-informative priors for the remaining parameters. The standard noninformative prior is used for σ , $p(\sigma) \propto 1/\sigma$. This prior is used despite the fact that its rationale as a Jeffreys prior no longer holds, given the form of the likelihood function in (3). Since this same prior for σ is used in all of M_1 to M_4 , the Bayes Factors used for all pairs of these models are unaffected by the fact that this prior is improper. For the free parameters in the *GARCH*(1, 1) specifications, a uniform prior is specified over the region bounded by $0 < b < 1$, $0 < c < 1$ and $0 < b + c < 1$.

An inverted gamma prior is specified for ν , with $E(\nu) = 10$ and $var(\nu) = 600$. The prior

is calibrated so as to match the location of the posterior density for ν based on estimation of a Student t model for the returns data, but with the variance of the prior being several-fold larger than the variance of the returns posterior. A normal prior is specified for γ , with $E(\gamma) = 1.0$ and $var(\gamma) = 0.08$. Again, the prior is calibrated to match the location of the posterior density for γ estimated from the returns data, but with the variance of the prior specified to be much larger.¹⁷

5.2.3 Results

In Table 2 we report the mean, mode and 95% HPD intervals for each parameter in the six models estimated. The first thing to note is the relative constancy, across the constant volatility models, of the point estimates of volatility. The modal estimate of σ varies only between 0.127 for M_1 and 0.130 for M_2 to M_4 . The mean estimates vary between 0.125 (M_2) and 0.130 (M_4). The interval estimates of σ for the larger models are, however, notably larger than that for M_1 , indicating that σ ranges between 0.121 and 0.139. The Black-Scholes model estimates σ within a very narrow band of 0.125 to 0.129.

The mean and modal point estimates of the degrees of freedom parameter, ν , in M_2 are both approximately equal to 7, with the 95% interval estimate covering values for ν which represent a distinct departure from normality in the returns process. The skewness parameter in M_4 , γ , is estimated to be less than unity, indicating negative skewness in the returns process. However, the 95% interval is very wide, easily covering values for γ which imply either symmetry ($\gamma = 1$) or positive skewness ($\gamma > 1$). These numerical results are illustrated graphically in Figure 1, in which the estimated marginal densities for models M_1 , M_2 and M_3 are presented.

The corresponding parameter estimates for the two *GARCH* models, M_5 and M_6 are very similar. The shock persistence indicated by the estimated *GARCH* models is quite large, with the sum of the point estimates of b and c for each model being approximately equal to 0.9.

¹⁷Experimentation has indicated that the implicit posterior densities for ν and γ produced using the returns-based priors differ only negligibly from posteriors produced from uniform priors. This is important, given that one point of interest is the contrast between the implicit and direct posteriors. However, the Bayes Factors related to the models in which ν and γ feature are well-defined only when proper priors are specified for these parameters. One way of avoiding the usual arbitrariness associated with the prior specification is to use the returns data to determine their essential form.

Table 2:
Implicit Marginal Posterior Parameter Densities

		Mode	Mean	95% HPDI
M_1	σ	0.127	0.127	(0.125, 0.129)
M_2	σ	0.130	0.125	(0.121, 0.139)
	ν	7.100	7.133	(6.100, 15.100)
M_3	σ	0.130	0.125	(0.121, 0.139)
	γ	0.700	0.746	(0.200, 1.300)
M_4	σ	0.130	0.130	(0.121, 0.139)
	ν	8.100	8.274	(5.100, 20.100)
	γ	0.750	0.750	Fixed
M_5	a	0.005	0.005	Fixed
	b	0.260	0.245	(0.235, 0.285)
	c	0.640	0.640	(0.610, 0.670)
M_6	a	0.005	0.005	Fixed
	b	0.220	0.204	(0.190, 0.250)
	c	0.650	0.650	(0.620, 0.680)
	ν	7.000	7.000	Fixed

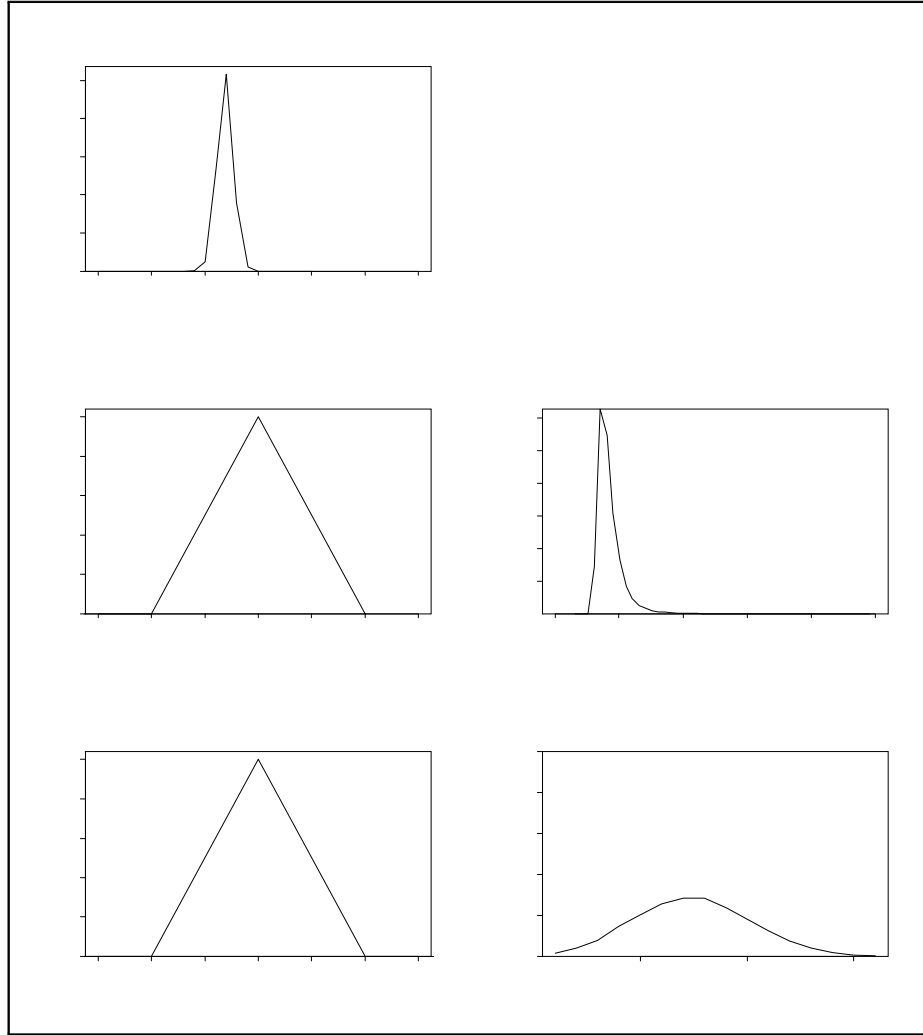


Figure 1: Implicit Marginal Posteriors for M_1 , M_2 and M_3 .

5.3 Model Rankings and Averaged Results

5.3.1 Implicit Bayes Factors and Model Probabilities

Panel A in Table 3 provides the estimated Bayes Factors for the four constant volatility models M_1 to M_4 , with the Black-Scholes model, M_1 , used as the reference model. The final row in Panel A gives the associated model probabilities, based on equal prior probabilities for all four models. The corresponding statistics for the time-varying volatility models, M_5 and M_6 , are given in Panel B, with M_5 used as the reference model in the construction of the Bayes Factors. The reason for considering these two sets of results separately is as follows. Since the *GARCH* models nest the corresponding constant volatility models, when $b = c = 0$ in (13), a Bayes Factor could, in principle, be applied to assess the evidence in the data for the *GARCH* models, relative to the constant volatility models, so long as a prior is placed on a in (13) which is consistent with the

Table 3:
 Implicit Bayes Factors and Model Probabilities
 Entry (i, j) indicates the Bayes Factor
 in favour of M_j versus M_i

Panel A					Panel B		
(Constant volatility models: M_1 to M_4)					(<i>GARCH</i> models: M_5 and M_6)		
	M_1	M_2	M_3	M_4		M_5	M_6
M_1	1.000	0.287	0.002	0.219	M_5	1.000	1.2143×10^{-13}
M_2		1.000	0.007	0.763	M_6		1.000
M_3			1.000	109.5			
M_4				1.000			
Prob(M_k)	0.663	0.190	0.001	0.145	Prob(M_k)	1.000	1.2143×10^{-13}

prior placed on σ in the constant volatility models. However, in our implementation of the implicit estimation method, we have restricted the parameter a in all of the *GARCH* models to be equal to 0.005, thereby assigning a degenerate prior to a . As such, it is not legitimate to construct Bayes factors for the *GARCH* models using any of the constant volatility models as a reference model since, in the latter case, σ is always a free parameter, with a non-degenerate prior.

With regard to the interpretation of the Bayes Factors and model probabilities within the separate constant volatility/*GARCH* blocks, it needs to be remembered that certain models are estimated in restricted form. For instance, all Bayes Factors in Panel A of Table 3 which involve M_4 measure the relative support in the data for a *specific* skewed Student t model for returns, namely one in which the skewness coefficient is set at $\gamma = 0.75$, the mean estimate of γ in M_3 . In the estimation of the marginal likelihood for M_4 , using the expression in (19), γ is not included as a free parameter in Θ_4 , but, rather, is used to characterize the model on which the evaluation of the likelihood, prior and posterior is conditional. The probability assigned to M_4 in Table 3 is the posterior probability of this specific skewed Student t model. Similarly, in Panel B of Table 3, the results need to be interpreted in the light of the fact that M_6 specifies a *GARCH* process allied with a *specific* Student t distribution, with $\nu = 7$.¹⁸

The results in Panel A of Table 3 indicate that the Black-Scholes option pricing model, with a

¹⁸Note that *both* of the two *GARCH* models estimated restrict a to equal 0.005.

Table 4:
Proportion of fit intervals which cover zero

M_1	M_2	M_3	M_4	M_5	M_6
0.044	0.726	0.102	0.303	0.077	0.053

posterior probability of 0.663, is given a substantially larger posterior weight by the option prices than any of the alternative constant volatility models. That said, the Student t (M_2) and skewed Student t (M_4) models get reasonable posterior weight, at 0.190 and 0.145 respectively. This result is consistent with the low value estimated for the degrees of freedom parameter, ν . Only the skewed normal model (M_3) gets negligible posterior weight, with a posterior probability of 0.001, this result tallying with the rather weak evidence in favour of skewness reported in Table 2. The Bayes Factors reported in Panel B for the *GARCH* models, imply that the Normal-*GARCH* (M_5) and Student- t -*GARCH* (M_6) models have posterior probabilities of one and zero respectively. This result can be interpreted as indicating that the time-varying volatility, on its own, captures all of the option-implied leptokurtosis in returns, without there being any need to specify, in addition, a nonnormal conditional distribution.

5.3.2 Out-of-Sample Fit

Table 4 summarizes the out-of-sample fit performance of the six different models. The figure reported for each model is the proportion of fit intervals as given in (21) which cover zero. As is evident, the more highly parameterized models have better out-of-sample fit performance than the single-parameter, normal-constant volatility model, M_1 . With no penalty for model size, this result is to be anticipated. The Student- t -constant volatility model, M_2 , provides by far the best out-of-sample fit of all the models considered. The fit proportions for the two-parameter, constant volatility models, M_2 , M_3 and M_4 , rank the models in the same way as the posterior probabilities reported in Table 3, namely: M_2 first, M_4 second, followed by M_3 . The fit proportions for the two *GARCH* models, M_5 and M_6 , also reflect the model probability rankings in Table 3, although not the extreme dominance of M_5 over M_6 that the model probabilities indicate.

5.3.3 Out-of-Sample Predictive Performance

The improvement in fit associated with an increase in the parameter set is obviated via the production of the predictive density, in which the parameters are (approximately) integrated out via (23).

The results in Table 5 thus enable an assessment of the predictive ability of the different models which is independent of the size of their parameter sets. It is clear that no one of the six models dominates overall in terms of predictive performance and that there is little to choose between the different models. The average probability ascribed to the bid-ask spread is approximately 22% to 23% for all models. Given that the bid-ask spreads are narrow relative to the spread of the predictive densities, these figures indicate that the predictives for all models tend to be reasonably well-centred, with the bid-ask spreads falling within the region of high predictive content. If the predictive mode or mean is used as a point predictor of the option price, the results in Table 5 show that the probability of predicting an option price within the observed spread ranges from 30% to 46%. Again, these figures are indicative of the predictives for all models being reasonably well-located. If the prediction criterion is taken as the full bid-ask spread lying within a high predictive probability interval, as measured by the interquartile range, M_2 does noticeably better than the other models, but with no model having a proportion greater than 29%. Once the prediction interval is expanded to have 95% probability content, essentially all bid-ask spreads are covered. The 13.8% which are excluded for all models correspond to intervals for which the bid price violates the no-arbitrage lower bound. If the predictive criterion relates to covering the market price (= the average of the bid and ask prices), then all models perform equally well, with *all* market prices being encompassed by the interquartile range of all model predictives (and hence by the 95% intervals also).

The final column in Table 5 contains the statistics associated with the averaged predictive density. Since the constant volatility and *GARCH* models need to be considered separately, and since the model probabilities for the *GARCH* models favour the normal-*GARCH* model so strongly over the Student t -*GARCH* model, we present the predictive density averaged only over the four constant volatility models, M_1 to M_4 . The model-averaged predictive is obtained by weighting the ordinates of the four predictives associated with models M_1 to M_4 by the model probabilities given at the bottom of Table 3. The averaged predictive is then used to calculate the various measures of predictive accuracy. Potentially, this process could produce results which are uniformly superior to the model specific results. In the present case, however, with all models performing in a similar manner, the averaging process produces results which are in themselves similar to the model specific results. Not surprisingly, given the large weight assigned to the Black-Scholes model (0.663), the averaged results are most similar to those of M_1 . Only in terms of the fourth predictive criterion, the proportion of bid-ask spreads in the interquartile range, would averaging across the four models produce a noticeably better result than would the choice of M_1 alone.

Table 5:
 Predictive Performance of the Different Models
 (All figures are proportions of the hold-out sample size of 413)

	M_1	M_2	M_3	M_4	M_5	M_6	Model Averaged (M_1 to M_4)
Prob(ba) ^(a)	0.232	0.222	0.226	0.228	0.232	0.233	0.230
Mode in ba	0.416	0.429	0.460	0.407	0.397	0.404	0.419
Mean in ba	0.310	0.308	0.300	0.305	0.317	0.322	0.308
ba in IQI ^(b)	0.179	0.286	0.169	0.186	0.165	0.172	0.191
ba in 95% I ^(c)	0.862	0.862	0.862	0.862	0.862	0.862	0.862
Price in IQI	1.000	1.000	1.000	1.000	1.000	0.993	1.000
Price in 95% I	1.000	1.000	1.000	1.000	1.000	1.000	1.000

(a) ba = the bid-ask spread

(b) IQI = the interquartile interval.

(c) 95% I = the interval which excludes 2.5% in the lower and upper tails of the predictive distribution. This interval equals the 95% HPD interval only for those predictives which are symmetric around a single mode. The 13.8% of bid-ask spreads which are excluded from the 95% intervals for all models, correspond to intervals for which the bid price violates the lower bound described in the text.

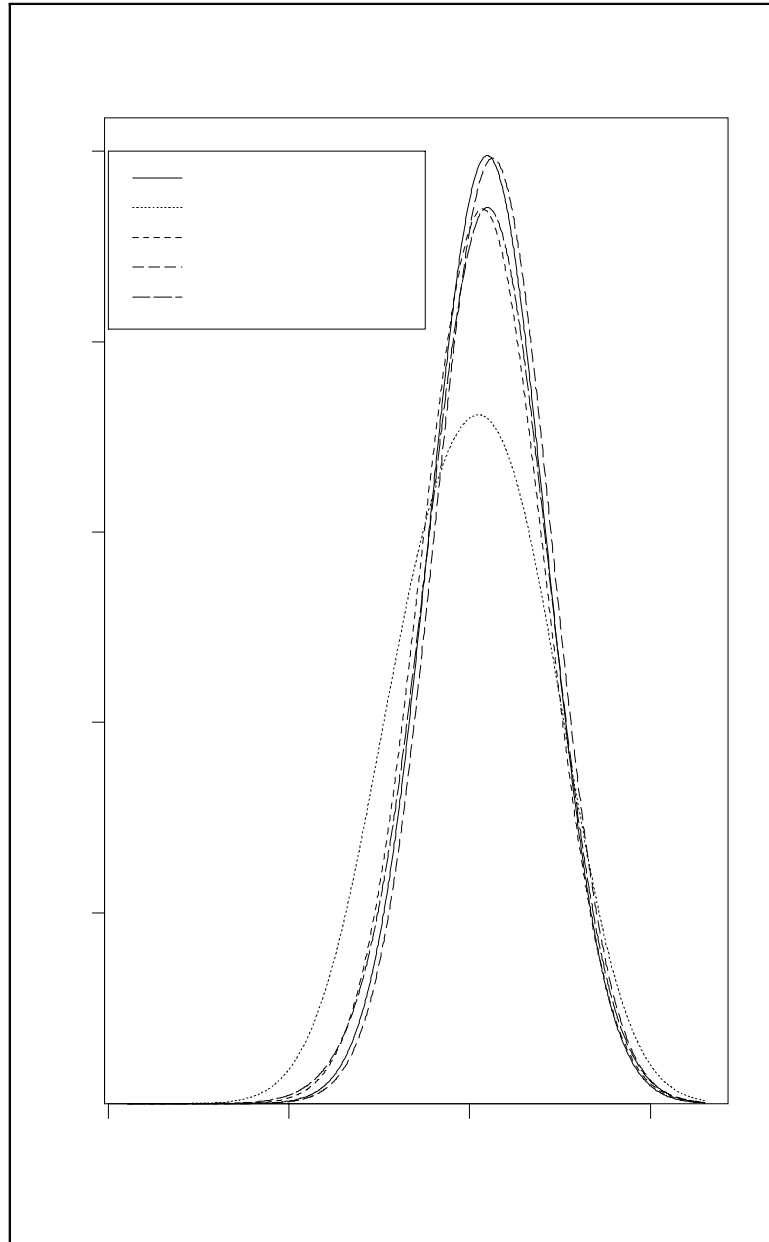


Figure 2: Predictive Densities for an *ATM* Contract

Figure 2 gives an example of the predictive densities associated with models M_1 , M_2 , M_3 and M_4 , for a specific *ATM* contract on the 29th of December ($S = 613.16$; $K = 615.00$; $S/X = 0.997$). The average bid-ask spread quoted for this contract during the day was (\$13.25 to \$13.82). The market price of this option contract is thus taken to be $c = \$13.53$. The averaged predictive density is included on the same graph. The Black-Scholes (M_1) and skewed Student t (M_4) predictives track one quite closely, with the latter being slightly to the right of the former. The modal values are \$15.49 and \$15.63 respectively. The skewed normal (M_3) predictive has slightly fatter tails than these two densities, as well as a smaller modal value of \$15.36. Only the Student t (M_2) predictive

for this particular option price is noticeably different, being more diffuse than the other densities. Its modal value of \$15.23 is the closest to the observed market price. Given the probability weight assigned to M_1 , the averaged predictive is not too dissimilar to that for the Black-Scholes model, with a modal value of \$15.45. However, the impact of M_2 has caused a slight leftward shift in the averaged density, relative to that of Black-Scholes, as well as rendering it more diffuse. As is clear from looking at a figure like Figure 2, if the different models were to have widely divergent predictives, the averaging process may well produce a predictive with much better coverage than that of any of the individual models.

5.3.4 Hedging Performance

The results of the hedging performance of the alternative option price models are given in Table 6. The hedge distributions are based on computing the delta hedge on the 28th of December, 1995, and evaluating the hedge error in (25) associated with the portfolio on the next day. That is, Δt in (25) equals $\Delta t = 1/365$. The calculations are performed on the beginning-of-day prices of 28 contracts, which is the number of contracts common to both trading days. For each simulated value of θ_k for model M_k , the average hedging error over all 28 contracts is calculated. The full set of B such average hedging errors is then used to estimate the posterior density of the hedging error, via kernel smoothing. This same process is followed for *OTM*, *ATM* and *ITM* contracts. The mean and approximate 95% HPD interval (in parentheses) of each hedging distribution is reported in Table 6. All figures are expressed in dollars.¹⁹ In the final column of Table 6, the modes and 95% HPD intervals of the model averaged hedging distributions are recorded, with the averaging occurring over the constant volatility models only. As is clear from the results in Table 6, the location of the hedging distributions is very similar across models. Only the variability differs across models, with the Student t (M_2) hedging distribution exhibiting the widest HPD intervals, for all contract groupings. The point and interval estimates reported in the table indicate that all models overhedge *ITM* and *ATM* contracts and underhedge *OTM* contracts. The averaging process produces hedging densities whose location reflects that of the model with the highest weight, namely M_1 . The variability of the averaged densities, as reflected in the width of the HPD intervals, is similar to that of the most variable model-specific densities, namely those of M_2 . Most significantly, none of the HPD intervals reported in Table 6 cover zero. This can be interpreted as meaning that all models considered are misspecified when it comes to hedging.²⁰ However, whether

¹⁹Some of the hedging densities are multimodal. However, we do not report discontinuous intervals, as would be the case for HPD intervals strictly defined. The multimodality is also the reason why we present mean hedging errors rather than modal hedging errors, for the model specific densities.

²⁰This result is somewhat consistent with that of Bakshi, Cao and Chen (1997), in which the hedging performance, based on a single instrument hedge, of the Black-Scholes and alternative stochastic volatility models was virtually

Table 6:
Hedging Performance of the Different Models (\$)

	M_1	M_2	M_3	M_4	M_5	M_6	Model Averaged (M_1 to M_4)
	Mean (HPD)	Mean (HPD)	Mean (HPD)	Mean (HPD)	Mean (HPD)	Mean (HPD)	Mode (HPD)
All	0.124 (0.123, 0.125)	0.125 (0.120, 0.128)	0.125 (0.123, 0.126)	0.125 (0.123, 0.126)	0.125 (0.124, 0.126)	0.125 (0.124, 0.126)	0.124 (0.122, 0.127)
<i>OTM</i>	-0.079 (-0.081, -0.077)	-0.078 (-0.086, -0.071)	-0.077 (-0.078, -0.076)	-0.079 (-0.082, -0.075)	-0.077 (-0.078, -0.076)	-0.078 (-0.080, -0.076)	-0.079 (-0.085, -0.072)
<i>ATM</i>	0.271 (0.270, 0.272)	0.271 (0.267, 0.274)	0.270 (0.269, 0.271)	0.270 (0.269, 0.272)	0.270 (0.269, 0.271)	0.270 (0.269, 0.271)	0.271 (0.268, 0.274)
<i>ITM</i>	0.137 (0.136, 0.138)	0.138 (0.136, 0.140)	0.139 (0.138, 0.140)	0.138 (0.136, 0.140)	0.139 (0.138, 0.140)	0.138 (0.137, 0.140)	0.137 (0.135, 0.141)

the observed hedging errors are significant from an economic point of view is unclear. The hedging errors range in magnitude from approximately 9 to 28 cents, whilst from Table 1 it can be seen that the option prices themselves range from an average price of about \$3.00 for *OTM* options to an average price of about \$80.00 for *ITM* options. Viewed in relation to the magnitude of the option prices, these hedging errors do not seem to be substantial. Figure 3 presents the averaged distributions for all four sets of contracts.

indistinguishable, apart from the case of *OTM* options. Our hedging results are based upon a substantially smaller sample of S&P500 option prices than are those of Bakshi, Cao and Chen.

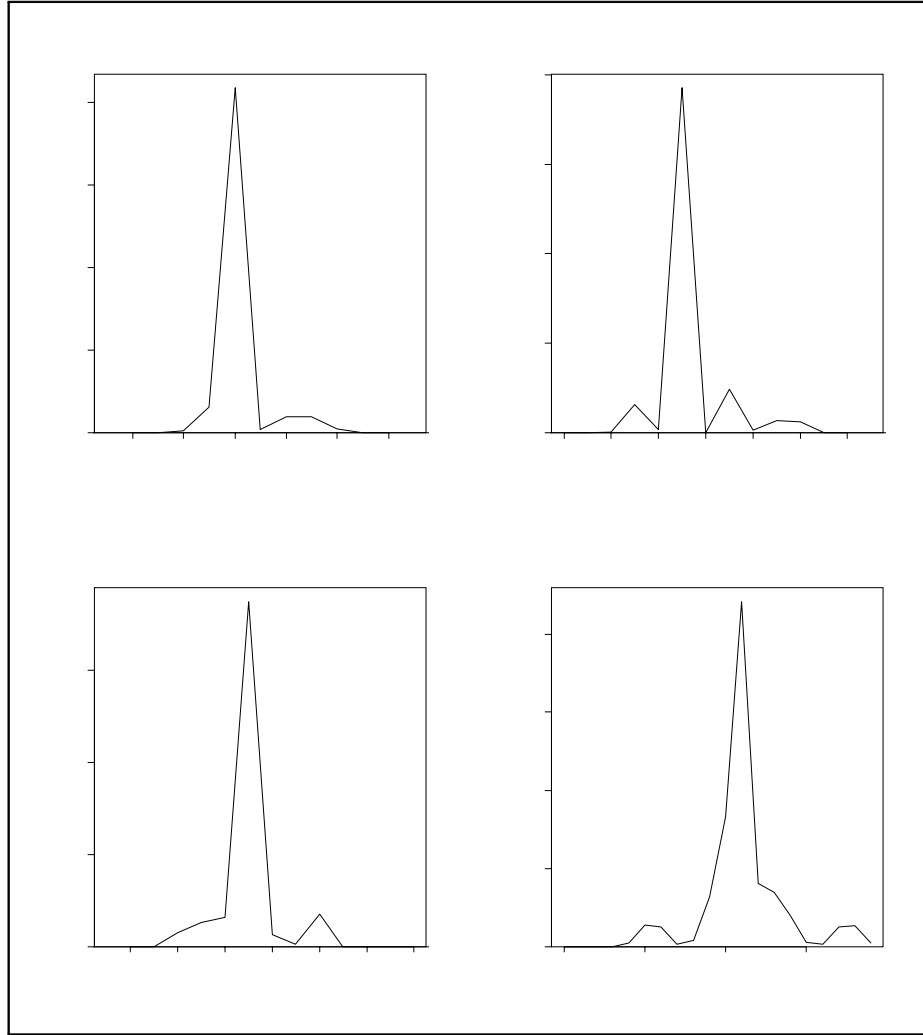


Figure 3: Averaged Hedging Posteriors (Error in Cents)

5.4 Comparison with Estimation Based on the Returns Data

We conclude the empirical analysis by reporting some results based on direct estimation of models M_1 to M_6 using returns data. The data used in the direct estimation is a sample of end-of-day day prices for the S&P500 stock index from the first trading day in 1995 up to December 27th, 1995. To account for possible serial correlation in the mean of (8), an autoregressive model of order two is fitted to the returns. With neither coefficient on lagged returns being significant, the unfiltered returns are used in the succeeding analysis.

Gibbs sampler algorithms are run for each model to produce estimates of the individual parameters and the Bayes factors.²¹ One thousand initial iterations are discarded and five thousand subsequent iterations used to estimate the joint posterior distribution. Marginal densities are cal-

²¹Griddy Gibbs samplers are used, except for in the case of M_1 , where the analytical full conditionals are used.

Table 7:
Returns-based Marginal Posterior Parameter Densities

		Mode	Mean	95% HPDI
M_1	$\sigma^{(a)}$	0.078	0.079	(0.072, 0.085)
M_2	$\sigma^{(a)}$	0.078	0.079	(0.070, 0.092)
	ν	5.000	8.110	(3.000, 19.000)
M_3	$\sigma^{(a)}$	0.079	0.079	(0.071, 0.086)
	γ	1.000	1.000	(0.856, 1.144)
M_4	$\sigma^{(a)}$	0.081	0.079	(0.070, 0.092)
	ν	5.000	7.780	(3.000, 20.000)
	γ	1.020	1.030	(0.880, 1.200)
M_5	a	0.010	0.100	(0.004, 0.224)
	b	0.030	0.050	(0.001, 0.140)
	c	0.920	0.570	(0.005, 0.970)
M_6	a	0.010	0.110	(0.001, 0.280)
	b	0.030	0.060	(0.001, 0.200)
	c	0.920	0.540	(0.005, 0.970)
	ν	5.000	7.380	(3000, 18.000)

(a) Since the returns data is daily data, scaled by a factor of 100, these estimates of σ have been scaled by a factor of $\sqrt{250}/100$, to render them consistent with the annualized figures which have been estimated from the option price data.

culated using Rao-Blackwellized estimators. The Bayes factors are calculated using the method of Chib (1995) based on five thousand additional iterations of suitably constrained Gibbs samplers. In both *GARCH* models, M_5 and M_6 , the parameters a and c are sampled from their bivariate full conditional distribution due to high negative correlation between the two parameters. Similarly, the parameters ν and σ in model M_4 are sampled jointly due to high correlation.

Estimates of the marginal posteriors are produced using either noninformative or proper, but diffuse, priors for all parameters. A-priori independence between all parameters is imposed. The standard noninformative priors are used for μ and σ , with $p(\mu) \propto 1$ and $p(\sigma) \propto 1/\sigma$ for models M_1 , M_2 , M_3 , and M_4 . The proper priors suggested by Fernandez and Steel (1998) for ν and γ are used, with $p(\nu) = 2 \exp(-\gamma^2/\pi)/\pi$, for $\gamma > 0$, and $p(\nu) = \exp(-\nu/10)/10$, for $\nu > 2$. For the *GARCH*

specification in M_5 and M_6 , an improper prior for a of the form $p(a) \propto 1/(2a)$ is adopted. Since this corresponds to the same prior adopted for σ^2 in the constant volatility models, the construction of Bayes Factors across the full set of models, M_1 to M_6 , is valid. For the remaining *GARCH* parameters in M_5 and M_6 , a uniform prior is specified over the region bounded by $0 < b < 1$, $0 < c < 1$ and $0 < b + c < 1$.

Table 7 presents the posterior parameter estimates, including point estimates based on the marginal mean and mode and 95% HPD interval estimates. The estimates of the constant volatility parameter are markedly smaller than the option-implied estimates, for all models M_1 to M_4 . The HPD intervals for σ also tend to be wider than those based on the options data. As is the case with the option-implied estimates, the introduction of additional distributional parameters in M_2 , M_3 and M_4 , produces point estimates of σ which differ very little, if at all, from the estimate for M_1 . Looking at the results for M_5 and M_6 , the evidence from the returns data on the persistence in conditional volatility is unclear, with the sum of the modal estimates of b and c being greater than that implied by the options data, at 0.95 for both M_5 and M_6 , but the sum based on the mean estimates being smaller, at 0.62 and 0.60 for M_5 and M_6 respectively.

It is interesting to note the degree of similarity between the estimates of ν and γ based on the returns data and the corresponding estimates based on the option price data. The estimates of ν in Table 7 range between 5.00 and 8.11, thereby being in the same range as the option-implied estimates in Table 2. The skewness point estimates in Table 7 are all close to one, with the 95% HPD interval covering values for γ which indicate all forms of skewness. Again, this tallies with the option-implied results in Table 2, although the returns-based inference on skewness is certainly much sharper than that produced from the option data.

The results in Table 8 indicate that the returns data strongly favours the Student t model-constant volatility (M_2) over all other models estimated, including the Black-Scholes model, M_1 , with the posterior probability of M_2 being 0.759. This is in marked contrast with the results in Table 3, in which the option price data strongly favour M_1 over all other constant volatility alternatives. Model M_4 gets some posterior probability, at 0.074. Models M_1 and M_3 are assigned probabilities of only 0.020 and 0.002 respectively. The very small weight assigned to the model which allows for skewness but not leptokurtosis, namely M_3 , is consistent with the result recorded for this model in Table 3. The normal-*GARCH* model, M_5 , also receives only small weight, with a probability of 0.004. Only when the *GARCH* specification is combined with a conditional Student- t specification in M_6 is the allowance for time-varying volatility given any credence, M_6 having a probability of 0.161. This favouring of M_6 over M_5 is the reverse of the option-based result.

In Table 9, we report the fit and predictive performance of the Black-Scholes option pricing

Table 8:
Returns-based Bayes Factors and Model Probabilities
(Constant volatility models: M_1 to M_4)

Entry (i, j) indicates the Bayes Factor
in favour of M_j versus M_i

	M_1	M_2	M_3	M_4	M_5	M_6
M_1	1	38.800	0.080	3.770	0.200	8.240
M_2		1	0.002	0.097	0.005	0.212
M_3			1	47.13	2.500	103.0
M_4				1	0.053	2.186
M_5					1	41.20
M_6						1
Prob(M_k)	0.020	0.759	0.002	0.074	0.004	0.161

model, estimated from the historical returns data. Results for the remaining models estimated from the returns data are similar to those in Table 9 and are therefore not reported.²² The contrast between certain of the results in Table 9 and the corresponding results in Tables 4, 5 and 6 is quite startling. Using returns data to estimate M_1 produces fit distributions for all 413 option prices in the hold-out sample which *all* fail to cover zero. The out-of-sample predictive performance of M_1 is also adversely affected by the use of the returns data to estimate the volatility parameter, σ . The prediction coverage statistics in Table 9 are all markedly smaller than the corresponding statistics in Table 5. In particular, the proportion of bid-ask spreads in the interquartile interval of the returns-based predictive is zero!

6 Conclusions

This paper has developed a Bayesian approach to the implicit estimation of returns models using option-price data. In contrast to existing classical work, the Bayesian method takes explicit account of both parameter and model uncertainty in option pricing. The paper differs from most other Bayesian work on option pricing, which has implemented returns-based estimation. It also represents a significant extension of the implicit methodology in Jacquier and Jarrow (2000), with a full set of alternative parametric models for returns able to be estimated and ranked using option-price

²²Hedging results for the returns-estimated models are also not reported, since they essentially provide the same evidence of the misspecification of the models as do the options-based results.

Table 9:
Fit and Predictive
Performance of M_1 Estimated from Returns

<u>Fit</u>	
Proportion of intervals covering zero	0.000
<u>Prediction</u>	
Prob(ba) ^(a)	0.157
Mode in ba	0.138
Mean in ba	0.196
ba in IQI ^(b)	0.000
ba in 95% I ^(c)	0.409
Price in IQI	0.976
Price in 95% I	1.000

data. The application of Bayesian model averaging to option pricing has also been demonstrated.

The results of applying the methodology to option price data on the S&P500 index have been interesting, with no one parametric model being ranked highest according to all criteria. The Black-Scholes model is assigned a very high posterior probability of being the correct model out of the four constant volatility models considered. The models in which excess kurtosis is allowed for are, however, still assigned non-negligible probability, with quite low estimates produced for the relevant degrees of freedom parameter. Once time-varying volatility is allowed for, the specification of a leptokurtic conditional distribution would appear to be unnecessary.

The Black-Scholes model holds its own in terms of out-of-sample prediction. However, the predictive performance of all models is very similar, with no one model being best according to all predictive criteria. The fit performance of the Black-Scholes model does not compare well with that of the other alternatives, especially the Student- t -constant volatility model, which has by far the best out-of-sample fit performance of all the models considered. In terms of hedging, all of the models appear to be equally misspecified, although the economic significance of the hedging errors observed is questionable.

With option prices being produced via the interaction of market participants invoking potentially different distributional assumptions in their pricing of options, model averaging has the

potential to produce predictive distributions for out-of-sample option prices with better coverage properties than those associated with any one individual model. However, for the sample of S&P500 option data considered in this paper, all models considered had similar predictive performance. Hence, the model averaging produced only marginal improvement. Similarly, the averaged hedging distributions were not markedly different from those associated with the individual models.

Finally, as is clear from a comparison of the option-based and returns-based estimation, the S&P500 option prices do not merely reflect the information contained in historical data on index returns. Most notably, despite the strong returns-based posterior probability evidence in favour of the models with a Student- t distributional specification, the implicit posterior probabilities still strongly favour the Black-Scholes model. No matter which model is used, however, estimation of the parameters of that model using option price data provides substantially better out-of-sample fit and predictive performance than does any model estimated using returns data.

References

- [1] Baillie, R.T. and Bollerslev, T. (1989), “The Message in daily Exchange Rates: A Conditional variance Tale” , *Journal of Business and Economic Statistics*, 7, 297-305.
- [2] Baillie, R.T. and Bollerslev, T. (1991), “Intra Day and Inter Market Volatility in Foreign Exchange Rates” , *Review of Economic Studies*, 58, 565-585.
- [3] Bakshi, G, Cao, C. and Chen, Z. (1997), “Empirical Performance of Alternative Option Pricing Models” , *Journal of Finance*, 52, 2003-2049.
- [4] Bauwens, L. and Lubrano, M. (1998), “Bayesian Inference on GARCH Models Using the Gibbs Sampler” , *Econometrics Journal*, 1, C23-C46.
- [5] Black, F. and Scholes, M. (1973), “The Pricing of Options and Corporate Liabilities” , *Journal of Political Economy*, 81, 637-659.
- [6] Bollerslev, T. (1986), “Generalized Autoregressive Conditional Heteroskedasticity” , *Journal of Econometrics*, 31, 307-328.
- [7] Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992) “ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence” , *Journal of Econometrics*, 52, 5-59.
- [8] Bollerslev, T. and Mikkelsen, H. (1999), “Long-term Equity Anticipation Securities and Stock Market Volatility Dynamics” , *Journal of Econometrics*, 92, 75-99.

- [9] Boyle, P.P. and Ananthanarayanan, A.L. (1977), "The Impact of Variance Estimation in Option Valuation Models", *Journal of Financial Economics*, 375-387.
- [10] Chib, S. (1995), "Marginal Likelihood from the Gibbs Output", *Journal of the American Statistical*, 90, 1313-1321.
- [11] Cho, Y-H and Engle, R.F. (1999), "Modeling the Impacts of Market Activity on Bi-ask Spreads in the Option Market", *NBER Working Paper Series*, No. 7331.
- [12] 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.
- [13] de Vries, C.G. (1994), "Stylized facts of Nominal Exchange Rate Returns", *The Handbook of International Economics*, van der Ploeg, F. (ed.), Balckweel, 348-389.
- [14] Dumas, B., Fleming, J. and Whaley, R. (1998), "Implied Volatility Functions: Empirical Tests," , *Journal of Finance*, 53, 2059-2106.
- [15] Engle, R.F. (1982), "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation", *Econometrica*, 50, 987-1008.
- [16] Engle, R.F. and Mustafa, C. (1992), "Implied ARCH Models From Option Prices", *Journal of Econometrics*, 52, 289-311.
- [17] Fernández, C. and Steel, M.F.J. (1998), "On Bayesian Modelling of Fat Tails and Skewness", *Journal of the American Statistical Association*, 93, 359-371.
- [18] Friedman, D. and Vandersteel, S. (1982), "Short-run Fluctuations in Foreign Exchange Rates: Evidence from the Data 1973-79", *Journal of International Economics*, 13, 171-186.
- [19] George, T.J. and Longstaff, F.A. (1993), "Bid-ask Spread and Trading Activity in the S&P 100 Index Options Market", *Journal of Financial and Quantitative Analysis*, 28, No. 3, 381-397.
- [20] Geweke, J., (1999), "Using Simulation Methods for Bayesian Econometric Models: Inference, Development and Communication", *Econometric Reviews*, 18, 1-73.
- [21] Gouriéroux, C. and Monfort, A. (1994), *Simulation Based Econometric Methods*, CORE Discussion Paper.
- [22] Harvey, A.C., Ruiz, E. and Sentana, E. (1992), "Unobserved Component Time Series Models with ARCH Disturbances", *Journal of Econometrics*, 52, 129-158.

- [23] Harvey, A.C., Ruiz, E. and Shephard, N. (1994), “Multivariate Stochastic Variance Models”, *Review of Economic Studies*, 61, 247-264.
- [24] Hull, J.C. (2000), *Options, Futures, and Other Derivative Securities*, 4rd ed., Prentice Hall, New Jersey.
- [25] Hull, J.C. and White, A. (1987), “The Pricing of Options on Assets with Stochastic Volatilities”, *Journal of Finance*, 42, 281-300.
- [26] Jacquier, E. and Jarrow, R. (2000), “Bayesian Analysis of Contingent Claim Model Error”, *Journal of Econometrics*, 94, 145-180.
- [27] Jarrow, R. and Rudd, A. (1982), “Approximate Option Valuation for Arbitrary Stochastic Processes”, *Journal of Financial Economics*, 10, 347-369.
- [28] Johnson, H. and Shanno, D. (1987), “Option Pricing when the Variance is Changing”, *Journal of Financial and Quantitative Analysis*, 22, 143-151.
- [29] Kim, S., Shephard, N. and Chib, C. (1998), “Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models”, *Review of Economic Studies*, forthcoming.
- [30] Korolyi, G. A. (1993), “A Bayesian Approach to Modelling Stock Return Volatility for Option Valuation”, *Journal of Financial and Quantitative Analysis*, 579-594.
- [31] 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.
- [32] Lim, G.C., Martin, G.M. and Martin, V.L. (2000), “Option Pricing with Generalized Distributional Forms”, Draft Paper.
- [33] Lye, J.,N., Martin, V.L. (1993), “Robust Estimation, Nonnormalities and Generalized Exponential Distributions”, *Journal of the American Statistical Association*, 88, 253-259.
- [34] 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.
- [35] Raftery, A.E., Madigan, D. and Hoeting, J.A., (1997), “Bayesian Model Averaging for Linear Regression Models”, *Journal of the American Statistical Association*, 92, 179-191.
- [36] Rubinstein, M. (1994), “Implied Binomial Trees”, *Journal of Finance*, 49, 771-818.

- [37] 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.
- [38] Scott, L.O. (1987), “Option Pricing when the Variance Changes Randomly: Theory, Estimation and an Application”, *Journal of Financial Quantitative Analysis*, 22, 419-438.
- [39] Wiggins, J.B. (1987), “Option Values Under Stochastic Volatility: Theory and Empirical Estimates”, *Journal of Financial Economics*, 19, 351-372.