Journal of Business & Economic Statistics

ISSN: 0735-0015 (Print) 1537-2707 (Online) Journal homepage: https://www.tandfonline.com/loi/ubes20

Estimating the Parameters of Stochastic Volatility Models Using Option Price Data

A. S. Hurn, K. A. Lindsay & A. J. McClelland

To cite this article: A. S. Hurn, K. A. Lindsay & A. J. McClelland (2015) Estimating the Parameters of Stochastic Volatility Models Using Option Price Data, Journal of Business & Economic Statistics, 33:4, 579-594, DOI: 10.1080/07350015.2014.981634

To link to this article: https://doi.org/10.1080/07350015.2014.981634

Estimating the Parameters of Stochastic Volatility Models Using Option Price Data

A. S. Hurn and K. A. Lindsay
School of Economics and Finance, Queensland University of Technology, Brisbane QLD 4000, Australia (s.hurn@qut.edu.au; kenneth.lindsay@glasgow.ac.uk)

A. J. McClelland
Sydney Numerix, Sydney NSW 2000, Australia (clelland@numerix.com)

This article describes a maximum likelihood method for estimating the parameters of the standard square-root stochastic volatility model and a variant of the model that includes jumps in equity prices. The model is fitted to data on the S&P 500 Index and the prices of vanilla options written on the index, for the period 1990 to 2011. The method is able to estimate both the parameters of the physical measure (associated with the index) and the parameters of the risk-neutral measure (associated with the options), including the volatility and jump risk premia. The estimation is implemented using a particle filter whose efficacy is demonstrated under simulation. The computational load of this estimation method, which previously has been prohibitive, is managed by the effective use of parallel computing using graphics processing units (GPUs). The empirical results indicate that the parameters of the models are reliably estimated and consistent with values reported in previous work. In particular, both the volatility risk premium and the jump risk premium are found to be significant.

KEY WORDS: Jumps; Maximum likelihood; Particle filter; Risk premia; Stochastic volatility.

1. INTRODUCTION

This article describes and implements a maximum likelihood method for estimating the parameters of an equity index model with latent stochastic volatility, and possibly jumps, by means of a particle filter. The fundamental advantage of this approach is that all the parameters of the model, including those characterizing the equity premium and the volatility premium, can be reliably identified in a way that maintains the internal consistency of the physical and risk-neutral measures. The estimation procedure is described and applied in the context of two stochastic volatility models, namely, the standard square-root stochastic volatility model due to Heston (1993) and an extension of this model to include jumps in equity prices due to Bates (1996). Notwithstanding the known shortcomings of these models (e.g., the Heston model does not include jump risk of any sort, while the Bates model includes jumps in equity prices but not in volatility), they are still used as benchmarks in many practical instances and consequently have received considerable attention in the literature (see e.g., Bakshi, Cao, and Chen 1997; Eraker, Johannes, and Polson 2003; Eraker 2004). It should be stressed at the outset, however, that these models are used only as specific examples to allow the econometric methodology to be fully developed. The proposed method itself is, however, not limited to any particular model and the extension to more involved models is a matter of detail alone, requiring no further significant conceptual development.

Time-series observations on equity prices provide information about the physical dynamics of the model. On the other hand, option prices provide a rich source of data for estimating the parameters of continuous-time financial models because they provide information about the behavior of the model in a risk-neutral environment. This observation provides a strong motivation for incorporating as many option prices as possible into the estimation of financial models and this has been the subject of considerable research effort. Estimation based on variants of the method of moments includes studies by Chernov and Ghysels (2000) and Pan (2002). Maximum likelihood estimation is used by Aït-Sahalia and Kimmel (2007) and Santa-Clara and Yan (2010). Markov chain Monte Carlo methods are used by Jones (2003), Eraker (2004), and Forbes, Martin, and Wright (2007). A drawback of many of these studies in which the physical model and the risk-neutral model are estimated simultaneously is that, typically, only a small number (between one and three) near-the-money options are used in the estimation. Larger numbers of options are used by Bates (2000), Broadie, Chernov, and Johannes (2007), Christoffersen, Jacobs, and Mimouni (2010), and Andersen, Fusari, and Todorov (2012).

Notwithstanding these recent advances, there remains significant scope for developing a comprehensive estimation method that simultaneously uses both time-series observations on an equity index and option pricing data. The fundamental contribution of this article is to demonstrate that parallel computing using GPUs makes estimation of stochastic volatility models (with or without jumps) by maximum likelihood using a particle filter a feasible proposition. In this filter, each particle is a single realization of the unobserved state(s) (volatility and intensity when jumps are present) and integration over the unobserved states is achieved by a Monte Carlo method. The filter described here is closest in spirit to those used by Johannes, Polson, and Stroud (2009) and Christoffersen, Jacobs, and Mimouni (2010). The particle filter is implemented in this article in such a way as to allow rapid evaluation of the likelihood

Journal of Business & Economic Statistics
October 2015, Vol. 33, No. 4
DOI: 10.1080/07350015.2014.981634

function. Not only can the parameters of both the physical model and the risk-neutral model be estimated simultaneously, but simulation studies designed to explore the efficacy of the particle filter are also possible within a reasonable time frame. The results presented here indicate that particle filtering is a powerful estimation method that offers substantial advances over existing methods, particularly when options prices are used.

The computational complexity of the particle filter is driven by the requirement that each particle must be used to price all the required options. Consequently each function evaluation of a search algorithm involves the calculation of around 50,000 model option prices per particle. On a high-quality desktop PC a realistic outcome is one function evaluation per day, which implies one full model estimation every 6 months thereby rendering estimation problematic and simulation infeasible. The innovations that allow this kind of computation to be reduced to a matter of hours as opposed to months stem from the fact that the particles are independent draws from the distribution of the latent variable(s) which means that a particle filtering algorithm lends itself to parallel computation. Indeed, in the parlance of computer theorists the problem is "embarrassingly parallel" and computation in this kind of environment is becoming increasingly of interest to econometricians (Durham and Geweke 2011, 2013). In this work, the computational burden of each filtering step is dealt with by parallelization on a graphics processing unit (GPU). The calculations were performed on two supercomputers both running CUDA C version 5.

It is well known that the likelihood function returned by a particle filter is not a smooth function of the parameters, a problem that makes estimation by maximum likelihood problematic (see, e.g., Pitt 2002 and the references therein). Note, however, that this problem is merely a matter of degree. It may be argued that any numerical procedure that relies on finite precision arithmetic returns a discontinuous function although the lack of continuity for the most part is masked by visible accuracy. The issue really only relates to the level of precision at which this effect becomes a concern. The problem here is ameliorated by deliberately choosing a very large number of particles so that, in practice, near the minimum the likelihood function is well resolved to about six decimal places of accuracy. Note that although computation of the Hessian at the optimum is problematic (due to both the dimension of the parameter space and the issue of precision), the gradient is relatively well determined. Standard errors based on the outer product of the gradient approximation to the covariance matrix appear to be consistent with the values achieved under simulation. Consequently the contention is that the discontinuity of the likelihood function is not serious enough to warrant any additional measures to enhance smoothness.

The estimation of the stochastic volatility models uses time-series data on index returns and the associated cross-section of option prices written on the index for each trading day. Specifically the dataset comprises the S&P 500 index from January 2, 1990, to June 30, 2012 (5672 trading days) and as many as 55,396 option prices (approximately 10 per day), computed as the average of the close-of-day bid and ask prices, with maturity of no more than 90 trading days and which also satisfy a set of consistency and no-arbitrage conditions.

2. MODELS OF STOCHASTIC VOLATILITY

2.1 The Benchmark Model

Let $W_1(t)$ and $W_2(t)$ be Wiener processes defined under the physical measure, $\mathbb{P}$, with respective differentials dW_1 and dW_2. The prototypical stochastic volatility model (SV) proposed by Heston (1993), assuming a linear equity premium, is

$$\begin{aligned} \frac{dS}{S} &= \left(r - q + \xi_s V\right) dt + \sqrt{V}\, dW_1, \\ dV &= \kappa(\gamma - V)\, dt + \sigma\sqrt{V}\, dW_2\,, \end{aligned} \tag{1}$$

where S is the equity index level, V is the volatility of the index, r is the risk-free rate of interest, q is the dividend yield, ξ_s is a constant parameter governing the equity premium, σ governs the diffusion of volatility, and $\mathbb{E}[dW_1 dW_2] = \rho\, dt$ so that ρ represents the instantaneous correlation between innovations in returns and volatility. The quantities κ and γ denote, respectively, the rate of mean reversion of volatility and the long-run level of volatility associated with the physical measure, $\mathbb{P}$. Written informally, the expected excess return on the index under the physical measure satisfies

$$\mathbb{E}\left[\frac{dS}{S}\right] - (r - q)\, dt = \xi_s\, V\, dt,$$

so that $\xi_s\, V$ is a premium earned by the asset in excess of the risk-free rate of interest net of dividends, namely, the equity premium.

The dynamics of the equity index and its volatility under the risk-neutral measure, $\mathbb{Q}$, are

$$\begin{aligned} \frac{dS}{S} &= \left(r - q\right) dt + \sqrt{V}\, d\widetilde{W}_1 \\ dV &= \kappa(\gamma - V)\, dt - \xi_v V\, dt + \sigma\sqrt{V}\, d\widetilde{W}_2\,, \end{aligned} \tag{2}$$

where $\xi_v V$ is the volatility premium and $\widetilde{W}_1$ and $\widetilde{W}_2$ are Wiener processes now defined under the measure $\mathbb{Q}$ with the same correlation coefficient as before, namely, ρ. While the presence of the volatility premium, $\xi_v V$, is required for the change from the physical to the risk-neutral measure, it also reflects the fundamental underlying principle that risk averse investors are willing to pay a premium to hedge against increases in volatility, which in turn requires that $\xi_v < 0$. The volatility equation in (2) is usually written as

$$dV = \widetilde{\kappa}(\widetilde{\gamma} - V)\, dt + \sigma\sqrt{V}\, d\widetilde{W}_2\,, \tag{3}$$

where

$$\widetilde{\kappa} = \kappa + \xi_v\,, \qquad \widetilde{\gamma} = \frac{\kappa\gamma}{\widetilde{\kappa}}\,. \tag{4}$$

Given the expectation of $\xi_v < 0$, it follows that $\widetilde{\gamma} > \gamma$, so that the mean value of volatility under the risk-neutral measure $\mathbb{Q}$ is expected to be higher than mean volatility under the physical measure $\mathbb{P}$.

While the value of ξ_s can be estimated from asset prices alone, compensation for volatility risk, ξ_v, cannot because it does not appear in the dynamics of the equity price under the physical measure. Option contracts, however, are priced under the risk-neutral measure and therefore observations of option prices provide a means of identifying ξ_v.

2.2 Jumps in Equity Prices

As previously let $W_1(t)$ and $W_2(t)$ be $\mathbb{P}$-measure Wiener processes with respective differentials dW_1 and dW_2. The Bates (1996) model (SVJ) generalizes the SV model by allowing the asset price to experience jumps with constant intensity λ. Using notation consistent with the SV model (1), the SVJ model under the physical measure is

$$\begin{aligned} \frac{dS}{S} &= \big(r - q + \xi_s V + \xi_J - \lambda\,\mathbb{E}[J-1]\big)\,dt \\ &\quad + \sqrt{V}\,dW_1 + (J-1)\,dN \\ dV &= \kappa(\gamma - V)\,dt + \sigma\sqrt{V}\,dW_2\,, \end{aligned} \tag{5}$$

where again $\mathbb{E}[dW_1 dW_2] = \rho\,dt$. In the SVJ model, dN is the differential of the counting process of jump events, J is a nonnegative random variable describing the ratio of the asset price immediately after a jump to that immediately before the jump, and ξ_J governs the compensation earned by investors for accepting exposure to jump risk. In theory, the compensation for jump risk has two components, namely, a timing component associated with the intensity of the jump process and a size component associated with the magnitude of the jumps. The jump premium can be written as

$$\xi_J = \lambda\,\mathbb{E}[J-1] - \widetilde{\lambda}\,\widetilde{\mathbb{E}}[J-1]\,,$$

where $\widetilde{\lambda}$ and $\widetilde{\mathbb{E}}[J-1]$ are the values of the intensity and the mean jump return under the $\mathbb{Q}$-measure. When this specification for ξ_J is incorporated into Equation (5), the final model for jumps in asset price under the $\mathbb{P}$-measure becomes

$$\begin{aligned} \frac{dS}{S} &= \big(r - q + \xi_s V - \widetilde{\lambda}\,\widetilde{\mathbb{E}}[J-1]\big)\,dt \\ &\quad + \sqrt{V}\,dW_1 + (J-1)\,dN \\ dV &= \kappa(\gamma - V)\,dt + \sigma\sqrt{V}\,dW_2\,. \end{aligned} \tag{6}$$

Under the risk-neutral measure Equation (6) becomes

$$\begin{aligned} \frac{dS}{S} &= \big(r - q - \widetilde{\lambda}\,\widetilde{\mathbb{E}}[J-1]\big)\,dt + \sqrt{V}\,d\widetilde{W}_1 + (J-1)\,d\widetilde{N}, \\ dV &= \widetilde{\kappa}\big(\widetilde{\gamma} - V\big)dt + \sigma\sqrt{V}\,d\widetilde{W}_2\,, \end{aligned} \tag{7}$$

where $\widetilde{\kappa}$ and $\widetilde{\gamma}$ are defined in terms of κ, γ, and ξ_v in exactly the same way as that given in (4). In implementing the SVJ model outlined in Equations (5) and (6), it will be assumed that $\log J \sim \mathrm{N}(\delta_J, \sigma_J^2)$ under both the $\mathbb{P}$- and $\mathbb{Q}$-measures, a specification that attributes the total jump risk premium to the timing component. This representation of $\log J$ has the property that $\mathbb{E}[J] = e^{\delta_J + \sigma_J^2/2}$, from which it follows that

$$\xi_J = (\lambda - \widetilde{\lambda})\big(\exp\big(\delta_J + \sigma_J^2/2\big) - 1\big)\,. \tag{8}$$

This specification of the jump risk premium is slightly different to that favored by Pan (2002), who preferred setting the timing premium to zero and focusing on the size component of the premium. Our experimentation suggested that specifying the jump premium in terms of timing is a more robust approach.

3. DEVELOPMENT OF THE LIKELIHOOD FRAMEWORK

This section makes the simplifying assumption that both the index, S, and the volatility, V, are observed. Of course this assumption is unrealistic because, in reality, volatility is a latent variable. The fact that volatility is unobserved is fully dealt with in Section 4 after the conceptual groundwork has been set out.

3.1 The Likelihood Function in the Absence of Options Data

Let $X_0, \ldots, X_T$ be observations of a system at discrete times $t_0, \ldots, t_T$, respectively, and let $Z_k = \{X_j\}_{j=0}^{j=k}$ denote the full history of observed data up to and including the time t_k. Consider first the case in which $X_k = (S_k, V_k)$, where S_k and V_k are, respectively, the index and volatility at time t_k. The general parameter estimation problem may be represented succinctly in the form

$$\widehat{\theta} \equiv \arg\max_{\theta} \frac{1}{T}\left(\log f\big(Z_0\big) + \sum_{k=1}^{T} \log f\big(X_k \mid Z_{k-1}\big)\right). \tag{9}$$

For a fully observed Markov system, the history Z_{k-1} is captured by the state X_{k-1} and consequently $f(X_k \mid Z_{k-1})$ is just the value of the probability density function $f(X_k \mid X_{k-1})$. In particular, if $X_k = \{S_k, V_k\}$, then $f(X_k \mid X_{k-1})$ is the transitional probability density function $f(X_k, \Delta_k \mid X_{k-1}; \theta)$ of the process $X = \{S, V\}$ under the physical measure for a transition of duration Δ_k starting at X_{k-1} and ending at X_k. The maximum likelihood estimate of θ is

$$\widehat{\theta} = \arg\max_{\theta} \frac{1}{T} \sum_{k=1}^{T} \log f(X_k, \Delta_k \mid X_{k-1}; \theta) \tag{10}$$

in which no information has been associated with the initial state $X_0 = \{S_0, V_0\}$.

Although conceptually straightforward, the parameter estimation problem posed in (10) presents severe difficulties in practice. First and foremost among these is that f is not known in closed form even for simple multivariate affine problems, although the characteristic function of the density may be known in some cases. In all cases, a full inversion of the characteristic function to obtain the complete density is a nontrivial numerical operation. Aït-Sahalia and Kimmel (2007) provided a close approximation for $\log f$, which gives reasonable results.

3.2 Introducing Options Data

Suppose that X_k also includes observations of the prices $\{H_k^{(1)}, \ldots, H_k^{(M)}\}$ of M options so that $X_k = (S_k, V_k, H_k^{(1)}, \ldots, H_k^{(M)})$, where, as before, S_k and V_k are the index and volatility at time t_k. Let $\{\mathcal{H}_k^{(1)}, \ldots, \mathcal{H}_k^{(M)}\}$ be the respective model prices of these options at time t_k. Maintaining for the moment the unrealistic assumption that the model is fully observed, each option price provides independent information concerning the risk-neutral parameters, which is incorporated within the maximum likelihood framework through the specification

$$f\big(X_k \mid X_{k-1}\big) = f_p(X_k, \Delta_k \mid X_{k-1}; \theta) \prod_{j=1}^{M} g(H_k^{(j)} \mid \mathcal{H}_k^{(j)}; \theta),$$

where $f(X_k, \Delta_k \mid X_{k-1}; \theta)$ is the usual transitional probability density function of the physical model associated with the transition of duration Δ_k from $\{S_{k-1}, V_{k-1}\}$ to $\{S_k, V_k\}$, and $g(H_k^{(j)} \mid \mathcal{H}_k^{(j)}; \theta)$ is the density of the observed option price $H_k^{(j)}$ conditional on the corresponding model price $\mathcal{H}_k^{(j)}$.

The contribution of the term $\log f\big(Z_0\big)$ in Equation (9) to the likelihood function, which has been ignored in the absence of observations of option prices, now contributes to the log-likelihood function via the information contained in the observed option prices at t_0. When index value, volatility, and option prices are simultaneously observed, the maximum likelihood parameter estimates satisfy

$$\begin{aligned}\widehat{\theta} = \arg\max_{\theta} \frac{1}{T} \Bigg[& \sum_{j=1}^{M} \log g(H_0^{(j)} \mid \mathcal{H}_0^{(j)}; \theta) \\ & + \sum_{k=1}^{T} \Big(\log f(\mathbf{X}_k, \Delta_k \mid \mathbf{X}_{k-1}; \theta) \\ & + \sum_{j=1}^{M} \log g(H_k^{(j)} \mid \mathcal{H}_k^{(j)}; \theta) \Big) \Bigg], \end{aligned} \tag{11}$$

in which all the parameters of the model, θ, are identifiable, including the equity and volatility premia and the jump timing premium in the case of the SVJ model.

3.3 The Distribution of Option Pricing Errors

To complete the development of the estimation framework, the distribution of pricing errors for options must be specified. Technically, there is an over-identification problem, since conditional on all state variables and parameters, option prices are known. To circumvent the possibility of singularity, it is necessary to assume a pricing error in which model price differs from the observed price in some way. It is also possible to assume an error structure for implied volatilities as discussed in Christoffersen and Jacobs (2004), which has the same effect.

As noted by Huang and Wu (2004), the construction of the pricing error is a delicate but important issue. A likelihood framework is adopted here, in the sense that the density of the observed option $H_k^{(j)}$ conditional on the model price $\mathcal{H}_k^{(j)}$, denoted $g(H_k^{(j)} \mid \mathcal{H}_k^{(j)}; \theta)$ is specified. The nature of the error is either *additive* in which the difference between the observed and model option price is a zero mean Gaussian deviate, or *multiplicative* in the sense that it is the difference between the logarithms of the observed and model option prices that is a zero mean Gaussian deviate. There are at least four possible different distributions of pricing errors that can be sensibly specified. These are as follows.

Error Specification 1 (ES1). This is an additive pricing error in which the observed price H is Gaussian distributed about the model prices $\mathcal{H}$ with a standard deviation proportional to the model price, that is,

$$H \sim \frac{1}{\sqrt{2\pi}\,\alpha\mathcal{H}} \exp\left[-\frac{\big(H - \mathcal{H}\big)^2}{2\alpha^2\mathcal{H}^2} \right]. \tag{12}$$

In this specification, α is a nondimensional scaling factor.

Error Specification 2 (ES2). The pricing error remains additive but now the observed price H is Gaussian distributed about the model prices $\mathcal{H}$ with constant standard deviation α, that is,

$$H \sim \frac{1}{\sqrt{2\pi}\,\alpha} \exp\left[-\frac{\big(H - \mathcal{H}\big)^2}{2\alpha^2} \right]. \tag{13}$$

This is the formulation favored by Forbes, Martin, and Wright (2007) in which α now represents a dollar value.

Error Specification 3 (ES3). This is a multiplicative pricing error specification in which $\log(H/\mathcal{H})$ is Gaussian distributed with mean value zero and constant standard deviation α, that is,

$$\log\frac{H}{\mathcal{H}} \sim \frac{1}{\sqrt{2\pi}\,\alpha} \exp\left[-\frac{\big(\log(H/\mathcal{H})\big)^2}{2\alpha^2} \right]. \tag{14}$$

Error Specification 4 (ES4). This is again a multiplicative pricing error specification in which H is log-normally distributed with mean value $\mathcal{H}$ and parameter α, that is,

$$H \sim \frac{1}{\sqrt{2\pi}\,\alpha H} \exp\left[-\frac{\big(\log(H/\mathcal{H}) + \alpha^2/2\big)^2}{2\alpha^2} \right]. \tag{15}$$

The important difference between ES3 and ES4 is that the latter has the property that the expectation of the observed price is the model price.

The specifications ES2 and ES3 correspond broadly speaking to the \$MSE and %MSE loss functions investigated by Christoffersen and Jacobs (2004), although of course their measures make no distributional assumptions on the pricing errors. All four error specifications ES1–ES4 are defined by comparing the actual price of the option, H and the model-based price $\mathcal{H}$ and therefore do not explore pricing error defined in terms of implied volatilities. The performance of the estimation method under simulation will be investigated for each of these pricing error distributions.

3.4 Computing Option Prices

Finally, the calculation of option prices from the model under the risk-neutral measure is a computationally demanding task because closed-form expressions are not available for even the simplest stochastic volatility models. However, highly accurate approximations to European option prices can be computed for the affine class of models, of which the Heston (1993) and Bates (1996) models are examples. These approximations capitalize on the fact that, although the density function of the equity index under the model is not known, its Fourier transform is available.

Numerical algorithms to price options differ in how this information is used. The family of methods usefully labeled as quadrature methods, which include the methods of Carr and Madan (1999), Duffie, Pan, and Singleton (2000), and Lewis

(2001), recover option prices (or components thereof) via inversion integrals or related identities involving the Fourier transform of the density function of the equity index. The crux of the problem is that any such inversion integrals must be evaluated via numerical quadrature, and the choice of scheme determines the convergence properties of the approximation method.

By comparison, Fang and Osterlee (2008) developed an alternative approach which they call the cosine series approximation method. In this approach, the density function of the equity index is approximated using a cosine series expansion, and the payoff function is then integrated with respect to this approximating density. These integrals are available in closed form for simple payoff functions, such as those of the vanilla puts and calls used in the context of parameter estimation in this article. The crucial insight here is that coefficients of the cosine series expansion can be well approximated by evaluating the Fourier transform of the density function at appropriate frequencies.

The quality of the approximation for quadrature-based methods depends on the number of nodes in the quadrature, while for the cosine method the quality depends on the number of terms used in the cosine series expansion. Fang and Osterlee (2008) formally investigated the convergence properties of the cosine method, demonstrating that convergence is essentially exponential in the number of terms, whereas for quadrature-based methods, convergence is algebraic in the number of nodes when Newton–Coates algorithms are employed. They also conducted numerous simulation experiments to compare the cosine method with the quadrature methods of Carr and Madan (1999) and Lord et al. (2008), which is closely related to the method of Lewis (2001). The simulation evidence supports the theoretical analysis and suggests that to achieve the same level of accuracy, quadrature-based schemes require significantly more quadrature nodes than the cosine method needs terms in the expansion. When applied specifically to the Heston model, the conclusion is that the cosine method appears to be approximately a factor of 20 times faster than the Carr–Madan method for the same level of accuracy and, consequently, the cosine method is used in this article to compute all option prices. Cuda C code to implement the algorithm will be provided on request.

4. THE PARTICLE FILTER

The development of the likelihood framework in Section 3 assumes that the system is fully observed. In practice, however, volatility, V, is a latent variable. Consider now the case in which the system retains its Markov property but V is not observed and is removed from the observable series X. The observations Z_k provide information about the historical behavior of volatility, which in turn is informative with respect to both the current value of volatility, V_k, and the future evolution of index returns and option prices. The procedure by which information in Z_{k-1} is "folded" recursively into the transitional probability density function $f(X_k \mid Z_{k-1})$ in Equation (9) is called a filtering rule. The complexity of the model in combination with the fact that observed option prices are strongly nonlinear functions of the underlying state variables means that the associated filtering procedure must be implemented numerically.

In overview the recursive filtering procedure advancing the calculation of the log-likelihood function from t_{k-1} to t_k proceeds in two steps taking as its starting point the filtered probability density function $f(V_{k-1} \mid Z_{k-1})$ of the unobserved volatility at time t_{k-1}. The steps are as follows.

1. Incorporate the observation X_k at time t_k into the filtered probability density function $f(V_{k-1} \mid Z_{k-1})$ of unobserved volatility at time t_{k-1} and compute the contribution $f(X_k \mid Z_{k-1})$ to the likelihood function from the observed data X_k at time t_k. This objective is achieved by first computing $f(X_k, V_k \mid Z_{k-1})$ as the value of the integral
$$f\big(X_k, V_k \mid Z_{k-1}\big) = \int_{\mathcal{V}} f\big(X_k \mid V_k, V_{k-1}\big)\, f\big(V_k \mid V_{k-1}\big) \times f\big(V_{k-1} \mid Z_{k-1}\big)\, dV_{k-1}\,, \qquad (16)$$
where $\mathcal{V}$ is the sample space of volatility. Thereafter, the contribution to the log-likelihood function at time t_k is
$$f\big(X_k \mid Z_{k-1}\big) = \int_{\mathcal{V}} f\big(X_k, V_k \mid Z_{k-1}\big)\, dV_k\,. \qquad (17)$$
2. Use Bayes' Theorem to construct $f(V_k \mid Z_k)$, the updated filtered probability density function of the unobserved volatility at t_k. The filtered probability density function of unobserved volatility at t_k is given by
$$f\big(V_k \mid Z_k\big) = \frac{f\big(X_k, V_k \mid Z_{k-1}\big)}{f\big(X_k \mid Z_{k-1}\big)} \qquad (18)$$
thereby completing the second stage of the two-stage procedure.

Options data are introduced into the calculation through the recognition that the conditional probability density function $f(X_k \mid V_k, V_{k-1})$ appearing in integral (16) is the product of a contribution from the index value alone, namely, $f(S_k \mid V_k, V_{k-1})$, and option prices conditioned on the volatility V_k at time t_k. Specifically,
$$f(X_k \mid V_k, V_{k-1}) = f\big(S_k \mid V_k, V_{k-1}\big) \prod_{j=1}^{M} f\big(H_k^{(j)} \mid S_k, V_k\big)\,. \qquad (19)$$
Naturally, the density $f(H_k^{(j)} \mid S_k, V_k) = g(H_k^{(j)} \mid \mathcal{H}_k^{(j)})$ given our assumption that the pricing errors are independent of the state.

The integrals in Equations (16) to (18) must be approximated numerically, a task that is achieved using the particle filter. Setting aside for the moment the question of initialization, consider a cloud of P particles available at t_{k-1} in which each particle represents an independent draw from the sample space of volatility (or in general from the space of latent variables). It is assumed that the particle cloud is an efficient summary of the distribution of the latent variables taking account of the information Z_{k-1}. In effect, the particle cloud is regarded as an efficient summary of the probability density function $f(V_{k-1} \mid Z_{k-1})$.

The first step in the calculation of $f(X_k \mid Z_{k-1})$ is to use a numerical scheme, here the Euler scheme, to advance the solution of the physical model from the state $\{S_{k-1}, V_{k-1}\}$ at t_{k-1} in m small steps of size $\Delta t = (t_{k+1} - t_k)/m$, where $m = 4$ in this work. This procedure is repeated for each particle in the cloud, with each particle starting at the same index value S_{k-1} but a different value of V_{k-1}. This phase of the updating procedure generates a sample from the transitional density $f(V_k \mid V_{k-1})$. Importantly, the advancement of each particle is independent of all others, which is crucial for the parallelization of this step.

The next step in the computation of expression (16) calculates $f(X_k \mid V_k, V_{k-1})$ for each particle from the asset price S_k and option prices $H_k^{(1)}, \ldots, H_k^{(M)}$ using formula (19), that is,

$$f\big(X_k \mid V_k, V_{k-1}\big) = f(S_k \mid V_k, V_{k-1}) \prod_{j=1}^{M} g\big(H_k^{(j)} \mid \mathcal{H}_k^{(j)}\big),$$

where the explicit dependence of the probability density function on the parameters θ has temporarily been suppressed for notational convenience. The density $f(S_k \mid V_k, V_{k-1})$ is approximated using the method suggested by Pedersen (1995) and Brandt and Santa-Clara (2002). In this approach, the transitional probability density function $f(S_k \mid V_k, V_{k-1})$ is approximated by a Gaussian probability density function with mean and variance determined by the penultimate state of the integration procedure over a step of size Δt (see, e.g., Jensen and Poulsen 2002).

The contributions to the likelihood made by the options are calculated from expressions (12) to (15) in which the model option price is determined for each particle from the index, S_k, and the volatility of the particle, V_k. This is the most computationally demanding step of the filter. Note, however, that the prices corresponding to a given particle can be computed independently of any other particle and are therefore amenable to parallel computing. This step represents the crux of the computational efficiency of the algorithm.

Given θ, the value of $f(X_k \mid V_k, V_{k-1}; \theta)$ for the pth particle is denoted by $\mathcal{L}_p(\theta)$. The value of the integral (17) is then estimated from $\mathcal{L}_1(\theta), \ldots, \mathcal{L}_P(\theta)$ by Monte Carlo integration over the cloud of particles at time t_{k-1}, so that the contribution to the likelihood at t_k for the observation X_k is

$$f\big(X_k \mid Z_{k-1}; \theta\big) = \frac{1}{P} \sum_{k=1}^{P} \mathcal{L}_k(\theta).$$

The final step of the filter requires the resampling of the P particles with distribution $f(V_k \mid Z_k)$. The particle V_k is reselected with replacement with probability

$$w_p = \frac{\mathcal{L}_p(\theta)}{\sum_{k=1}^{P} \mathcal{L}_k(\theta)}.$$

Evaluating the contribution to the log-likelihood function at t_k and performing the resampling step involves all P particles. By comparison with the previous computations of the filter, the final step is not particularly intensive.

5. SIMULATION EXPERIMENTS

The efficacy of the particle filtering algorithm for estimating the parameters of stochastic volatility models is now examined in terms of a series of simulation experiments based on the SV model chosen because of its relative simplicity. The primary objective of the experiments is to gauge the efficacy of the particle filter algorithm in recovering the parameters of the stochastic volatility model and to investigate the various choices for the distribution of pricing errors in the option.

The first experiment compares the parameter estimates obtained by using the particle filter without any option prices with those obtained using discrete maximum likelihood (DML) based on a Gaussian approximation to the transitional density (Florens-Zmirou 1989; Shoji and Ozaki 1997; Elerian 1998) in which volatility is treated as observed. A slight refinement to the crude DML is used in that the mean and variance of the Gaussian distribution for each transition is actually the true mean and variance of the model (Hurn, Lindsay, and McClelland in press).

The primary objective of this simulation exercise is to obtain a benchmark bias for the estimates returned by the particle filter in the absence of option price data. Of course, the useful by-product is that the power of the particle filter is demonstrated in a situation in which the odds are stacked against it in terms of the alternative method. Despite the fact that the experiment is biased against the particle filter, it will be demonstrated that the DML approach based on the Gaussian density is inferior to the particle filter.

The experiment is set up as follows. Ten years of daily data ($T = 2520$) are simulated with $\Delta t = 1/252$ (so that the parameters may be interpreted as annualized values) and the parameter estimation problem is repeated 500 times. Importantly, 100 intermediate steps are used to simulate each daily increment to minimize discretization error in the simulated data series. The particle filter algorithm uses 17,920 particles, which was selected for its divisibility by 448, the number of cores on a Tesla 2070 card. For DML estimation, data are also generated for additional sample sizes of $T = 5040$ and $T = 7560$, again with $\Delta t = 1/252$.

The numerical optimization of the log-likelihood function constructed via the particle filter is performed using the downhill simplex method (Nelder and Mead 1965) as expounded in Gill, Murray, and Wright (1981). The downhill simplex method must be initialized not with a single starting guess but with an initial simplex that contains $n + 1$ vertices, where n is the dimension of the parameter space. In effect this requires the provision of $n + 1$ starting parameter vectors. The true parameters used in the simulation are used to generate the starting simplex by perturbing each parameter in such a way that it experiences a variability of 10% of absolute value across the simplex. Of course, in numerical optimization, there is always a danger of converging to a local optimum near the starting values. However, starting the optimization from an $(n + 1)$-dimensional simplex in each case does provide some protection against converging to local optima. More intensive strategies designed to alleviate the problem of local optima, such as using a grid of starting values, would be very difficult to implement in the current context because of computational considerations and the dimension of the parameter space.

The results of the simulation are reported in Table 1. The particle filter estimates of the drift parameter κ and the equity premium ξ_s for $T = 2520$ are significantly better than those provided by DML. Indeed, it is only when $T = 7560$ that the DML estimates match the quality of those of the particle filter. Recall that DML estimation treats volatility as observed whereas it is not treated as an observed variable by the particle filter. However, the particle filter introduces intermediate simulation steps when propagating particles and so provides a much better approximation to the true transitional density than the simple Gaussian approximation used by DML. This serves

Table 1. Parameter estimates for the SV model without option price data for $T = 2520$ with time step $\Delta t = 1/252$

Param.	True value	PF ($T = 2520$) Bias	PF ($T = 2520$) RMSE	DML ($T = 2520$) Bias	DML ($T = 2520$) RMSE	DML ($T = 5040$) Bias	DML ($T = 5040$) RMSE	DML ($T = 7560$) Bias	DML ($T = 7560$) RMSE
κ	3.00	0.1000	0.2203	0.4076	0.9312	0.1917	0.5640	0.1292	0.4466
γ	0.03	−0.0003	0.0037	−0.0007	0.0058	−0.0004	0.0038	−0.0002	0.0031
σ	0.30	−0.0010	0.0243	0.0003	0.0040	0.0001	0.0027	0.0001	0.0023
ρ	−0.60	0.0058	0.0341	−0.0001	0.0114	0.0007	0.0083	0.0010	0.0069
ξ_s	5.00	0.1727	0.5541	0.4881	1.9948	0.2238	1.3478	0.1221	1.1266

NOTE: The estimates from the particle filter (PF) using 17,920 particles for simulation experiments involving 500 repetitions are compared with discrete maximum likelihood estimation (DML) in which volatility is treated as an observed variable.

to explain the superior performance of the particle filter despite its informational disadvantage.

The next four simulation experiments investigate the performance of the particle filter when options data are used. To facilitate comparison, all the experiments were conducted on a Tesla 2070 card with pseudo daily data, $\Delta t = 1/252$, $T = 2520$, 17,920 particles and 6, 12, and 18 options per day. For each repetition, the initial value of volatility is set at the long-run mean and then burned-in for a period of 252 days using 24 Euler steps per day. Once the volatility has been burned in the simulation of the index starts with the value of 100.

In generating the simulated option prices, the path of the simulated price index and its volatility are identical for each set of options. For each repetition of the experiment, the option prices change only because of the choice of the distribution of pricing errors. Options are spread uniformly across the money when initiated. Three levels of maturity, 1 month, 2 months, and 3 months, are used and each option is followed through its lifetime until 5 days prior to maturity at which point the option is dropped from the set and a new option is generated. This generation procedure mimics the method proposed by Eraker (2004) for choosing options to use in estimation. Tables 2–5 report the results of the simulations for pricing-error distributions ES1 to ES4 in Equations (12) to (15). In each case, the synthetic option prices are generated using the same distribution that is used in the estimation to construct the likelihood.

The most important result to emerge from scrutiny of these results is that the introduction of even as few as six option prices into the estimation substantially reduces the bias on the parameters reported in Table 1. In addition, the volatility premium, ξ_v, is identified and accurately estimated. Interestingly enough, the volatility premium is recovered with very good precision but the estimate of the equity premium, ξ_s, is still somewhat imprecise. The reason for this is not entirely clear and more simulation evidence is required.

The pricing error introduced into the simulated prices, $\alpha = 5\%$, is recovered with high accuracy. Increasing the number of options from 6 to 18, by contrast, makes little significant difference to the root mean square errors particularly when interpreted in the context of the large amount of extra data that is processed. This is perhaps to be expected because the options are all priced using the correct model together with a moderate level of price noise, so that the addition of extra options adds little in terms of new information but significantly increases the amount of arithmetic to be done in the estimation process.

There is little to choose between the different distributions of pricing errors in terms of the quality of the parameter estimates. There are, however, significant differences in terms of the computational effort required to obtain these parameter estimates. Table 6 gives the time required for a single estimation of the various models implemented on one Xeon core hosting a single Tesla 2070 card.

Recall that ES2 is different from the others in the respect that the mispricing parameter α is independent of the model price and has the interpretation of a dollar value. This model, although relatively easy to estimate in terms of the results reported in Table 6, is perhaps the least realistic of the models when dealing with traded options. By contrast, ES1, ES3, and ES4 are concerned with the properties of the ratio of the observed to model price and within this class, and of these ES3 appears to require the least computational effort.

Table 2. Parameter estimates for the SV model with option pricing error ES1

Parameter	True value	6 Options Bias	6 Options RMSE	12 Options Bias	12 Options RMSE	18 Options Bias	18 Options RMSE
κ	3.00	0.0164	0.0339	0.0186	0.0313	0.0175	0.0272
γ	0.03	−0.0002	0.0003	−0.0002	0.0002	−0.0002	0.0002
σ	0.30	0.0007	0.0014	0.0009	0.0014	0.0007	0.0011
ρ	−0.60	0.0016	0.0031	0.0016	0.0026	0.0014	0.0021
ξ_s	5.00	0.4285	0.4761	0.4452	0.4799	0.4490	0.4823
ξ_v	−3.00	−0.0096	0.0108	−0.0097	0.0106	−0.0094	0.0103
α	0.05	0.0000	0.0005	0.0000	0.0004	0.0001	0.0004

NOTE: Estimation uses 17,920 particles for simulation experiments involving 500 repetitions with a fixed set of option prices.

Table 3. Parameter estimates for the SV model with option pricing error ES2

Parameter	True value	6 Options		12 Options		18 Options	
		Bias	RMSE	Bias	RMSE	Bias	RMSE
κ	3.00	0.0165	0.0298	0.0149	0.0295	0.0163	0.0299
γ	0.03	−0.0002	0.0003	−0.0001	0.0002	−0.0001	0.0002
σ	0.30	0.0007	0.0013	0.0006	0.0012	0.0008	0.0013
ρ	−0.60	0.0020	0.0031	0.0015	0.0029	0.0020	0.0030
ξ_s	5.00	0.4283	0.4405	0.4068	0.4234	0.3945	0.4073
ξ_v	−3.00	−0.0108	0.0111	−0.0100	0.0104	−0.0106	0.0109
α	0.05	0.0003	0.0008	0.0009	0.0012	0.0009	0.0011

NOTE: Estimation uses 17,920 particles for simulation experiments involving 500 repetitions with a fixed set of option prices.

Of course, the results of these experiments offer no real guidance as to the correct choice for the distribution of pricing error, but they do indicate that within a group of pricing error specifications that superficially appear to require comparable amounts of calculation and embody comparable accuracy as determined by the value of α (with the possible exception of ES2), the numerical effort in estimation is not necessarily similar.

Notwithstanding the computational effort required by ES4, in the absence of any evidence to the contrary, it would appear reasonable to conclude that using ES4 is a sensible choice because the mean value of the option price is not distorted by the choice of pricing error distribution. ES3 would be an attractive choice from a computational perspective, but the argument here is that the distortion of the mean price is undesirable and consequently only results for ES4 will be reported in Sections 7 and 8. Extensive experimentation revealed that the choice of pricing error distribution did not materially affect parameter estimation and results using ES3 are available on request.

6. OPTIONS DATA

The dataset used in the estimation to follow consists of daily S&P 500 index values from January 2, 1990, to June 30, 2012 together with the risk-free rate of interest and the dividend yield. The risk-free market factor used here is provided by Ken French, downloadable from his website *http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/*. The monthly dividend series used to construct the yield is downloaded from Robert Shiller's webpage *http://aida.wss.yale.edu/shiller/data.htm*. The options data used in the estimation comprise observations on bid and ask prices of all put and call options written on the S&P 500 index during the sample period, which are traded on the Chicago Board Options Exchange. The options data are obtained from *Delta-Neutral.com*, which is the provider visitors of *CBOE.com* are directed to when seeking historical data.

The complete sample of option price data amounted to 4,459,751 observations on options prices over 5672 trading days. Prior to estimation, prices of options are excluded from the set to be used in parameter estimation for days on which the option is not traded. To identify whether a contract is not traded on a given day, its daily trading volume is inspected. If this is recorded as zero within the dataset, this option is removed from consideration for selection. There are 3,165,523 instances of quotes excluded from the dataset in this manner. In addition, all options with negative bid-ask spreads are also excluded. The remaining options on each day are ranked in descending order of trading volume and subjected to two further tests.

1. *Bounds test*
   Suppose that C_{ask} and P_{ask} are, respectively, the ask prices for call and put options with strike price K and maturity T on an asset with spot price S. The *bounds-test* validates these prices against the theoretical bounds underlying the option, namely, that

$$\begin{aligned} C_{\text{ask}} &\geq \max(0, Se^{-qT} - Ke^{-rT}) \\ P_{\text{ask}} &\geq \max(0, Ke^{-rT} - Se^{-qT}), \end{aligned} \tag{20}$$

   where r is the risk-free rate during the period of the investment and q is the dividend yield. Failure to satisfy this

Table 4. Parameter estimates for the SV model with option pricing error ES3

Parameter	True value	6 Options		12 Options		18 Options	
		Bias	RMSE	Bias	RMSE	Bias	RMSE
κ	3.00	0.0302	0.0430	0.0225	0.0353	0.0200	0.0329
γ	0.03	−0.0003	0.0004	−0.0002	0.0003	−0.0002	0.0002
σ	0.30	0.0011	0.0016	0.0007	0.0014	0.0006	0.0012
ρ	−0.60	0.0022	0.0034	0.0011	0.0027	0.0011	0.0022
ξ_s	5.00	0.4108	0.4182	0.3974	0.4055	0.4027	0.4128
ξ_v	−3.00	−0.0104	0.0106	−0.0106	0.0108	−0.0106	0.0108
α	0.05	0.0000	0.0006	0.0000	0.0005	0.0001	0.0005

NOTE: Estimation uses 17,920 particles for simulation experiments involving 500 repetitions with a fixed set of option prices.

Table 5. Parameter estimates for the SV model with option pricing error ES4

Parameter	True value	6 Options Bias	6 Options RMSE	12 Options Bias	12 Options RMSE	18 Options Bias	18 Options RMSE
κ	3.00	0.0234	0.0400	0.0152	0.0288	0.0144	0.0263
γ	0.03	−0.0002	0.0003	−0.0002	0.0002	−0.0001	0.0002
σ	0.30	0.0009	0.0014	0.0006	0.0011	0.0005	0.0010
ρ	−0.60	0.0018	0.0029	0.0010	0.0021	0.0010	0.0018
ξ_s	5.00	0.4310	0.4514	0.4490	0.4828	0.4744	0.5067
ξ_v	−3.00	−0.0102	0.0107	−0.0097	0.0105	−0.0092	0.0100
α	0.05	0.0000	0.0006	0.0000	0.0004	−0.0001	0.0004

NOTE: Estimation uses 17,920 particles for simulation experiments involving 500 repetitions with a fixed set of option prices.

condition means that there is an opportunity for arbitrage by buying a single option and therefore options that fail the test are excluded.

2. *Arbitrage free test*
When call and put options with the same strike and maturity are traded on an asset, consistent bid and ask prices must satisfy the arbitrage conditions

$$\varphi(C_{\text{ask}} + Ke^{-rT}) \geq Se^{-qT} + P_{\text{bid}}$$
$$\varphi(Se^{-qT} + P_{\text{ask}}) \geq C_{\text{bid}} + Ke^{-rT}, \tag{21}$$

in which the right-hand side of the inequalities in (21) is the cash outflow and the left-hand side is the cash benefit. The parameter φ controls for transactions costs and is here set equal to 1.02. This is a generous allowance but is chosen to reflect the fact that end of day prices are being used.

The total number of options remaining after the application of these filters is 1,011,845 distributed over the 5672 trading days. The maturity of the options used in the estimation is then limited to be between 5 and 90 trading days leaving a total of 672,626 options potentially available for use in the estimation. For the remaining options, the average of their respective closing bid and ask prices on a given trading day are treated as their respective market-observed prices for that day when estimating parameters. The final dataset contains a variety of option quotes for different values of moneyness and maturity as shown in Table 7.

Table 6. The mean and standard deviations of the time taken (measured in seconds) for a single estimation of the SV model using 17,920 particles are given based on a sample of 500 trials of each combination of option count and distribution of pricing error

Error specification	Number of options: 0	6	12	18
ES1	1438	939	1640	2661
	(252)	(201)	(813)	(1259)
ES2	1438	931	1132	1944
	(252)	(210)	(183)	(428)
ES3	1438	939	1117	1945
	(252)	(201)	(145)	(229)
ES4	1438	1004	1707	2829
	(252)	(355)	(847)	(1315)

NOTE: Each estimation is performed using a Tesla 2070 card to facilitate comparison.

The final question to be resolved is how to choose from the remaining options exactly which of these should be used in the estimation. Eraker (2004) used a sampling strategy in which a contract that trades on the first day of the sample is chosen at random. All observations of this contract are then included in the estimation. On the first trading day on which this option ceases to trade, another trading contract is drawn at random. The sample of observed option prices is then constructed in this way so that there is at least one traded option each day. By contrast, the method for choosing options used here follows Pan (2002) who concentrated on the most liquid instruments, the implicit assumption being that the prices of these contracts should convey the most precise information. An additional benefit of following this approach is that it becomes more likely to select out-of-the-money options for inclusion, seeing as these are typically the most heavily traded contracts. As argued by Huang and Wu (2004) using out-of-the-money options for estimation and inference has become the industry standard for several reasons, one of which is that in-the-money options have positive intrinsic value which is insensitive to model specification.

As noted previously, the options on each day are ranked in descending order of trading volume and the requisite number of options for the estimation are then selected from this ordering, starting with the most liquid. Table 8 reports the distribution in terms of moneyness of the call and put options actually used in the estimation when 10 options per day are required. The total number of options used, 55,396, is slightly below the required value of 56,720 indicating that there are a number of days (mainly early in the sample period) where 10 options are not available. The distribution indicates that put options are slightly favored, a bias which is characteristic of the full sample both filtered, as shown in Table 7, and the full unfiltered dataset.

7. ESTIMATING THE SV MODEL

The parameters of the SV model are now estimated using both the index data, which is plotted in Figure 1, and the options data described in Section 6. Parameter estimates are presented for various combinations of options, namely, 1 in-the-money and 1 out-of-the-money option (1–1), 2 out-of-the-money options (0–2), 4 out-of-the-money options (0–4), 6 out-of-the-money options (0–6), and 10 out-of-the-money options (0–10). Each estimation is based on 35,840 particles.

Table 7. The number of European call and put options written on the S&P 500 Index that remain in the filtered dataset, tabulated by moneyness (index value over the option strike price) and maturity (days)

Range for spot to strike ratio	$5 \le T \le 22$ C	$5 \le T \le 22$ P	$22 < T \le 45$ C	$22 < T \le 45$ P	$45 < T \le 67$ C	$45 < T \le 67$ P	$67 < T \le 90$ C	$67 < T \le 90$ P	Row total
<0.90	7346	4103	14,278	3792	9156	2497	5525	1689	48,386
[0.90,0.92)	4827	1741	7918	1670	4489	859	1976	574	24,054
[0.92,0.94)	7522	2764	10,444	2635	5551	1267	2299	804	33,286
[0.94,0.96)	10,522	4773	13,109	4472	6376	2077	2489	1163	44,981
[0.96,0.98)	12,368	8194	13,552	7559	6401	3331	2411	1692	55,508
[0.98,1.00)	11,888	11,173	13,079	11,318	7177	6348	2563	2538	66,084
[1.00,1.02)	10,757	11,448	10,475	12,540	5450	7387	2012	294	63,010
[1.02,1.04)	8207	11,312	6724	11,884	2751	6359	1123	2707	51,067
[1.04,1.06)	5358	10,956	4356	11,196	1658	5707	786	2555	42,572
[1.06,1.08)	3417	10,111	2793	10,237	1017	4915	566	2342	35,398
[1.08,1.10)	2249	8602	1827	9027	710	4395	399	2118	29,327
> 1.10	6803	46,724	5755	61,259	3071	33,994	1577	19,441	178,624
Totals	91,264	131,901	104,310	147,589	53,807	79,136	23,726	40,564	672,626

Numbers of call and put options T trading days to maturity.

NOTE: The sample period is January 2, 1990 to June 30, 2012 comprising 5672 trading days.

Table 8. The distribution of European call and put options written on the S&P 500 Index that are actually used in the estimation of the stochastic volatility models when 10 out-of-the money options are required

Range for spot to strike ratio	$5 \le T \le 22$ C	$5 \le T \le 22$ P	$22 < T \le 45$ C	$22 < T \le 45$ P	$45 < T \le 67$ C	$45 < T \le 67$ P	$67 < T \le 90$ C	$67 < T \le 90$ P	Row total
< 0.90	16	0	31	0	16	0	1	0	64
[0.90,0.92)	19	0	56	0	25	0	7	0	107
[0.92,0.94)	99	0	108	0	86	0	22	0	315
[0.94,0.96)	293	0	497	0	292	0	111	0	1193
[0.96,0.98)	1082	0	1603	0	884	0	318	0	3887
[0.98,1.00)	6821	0	7845	0	4560	0	1589	0	20,815
[1.00,1.02)	0	6808	0	7834	0	4876	0	1914	21,432
[1.02,1.04)	0	1124	0	1796	0	1177	0	504	4601
[1.04,1.06)	0	443	0	636	0	409	0	165	1653
[1.06,1.08)	0	192	0	301	0	179	0	55	727
[1.08,1.10)	0	77	0	112	0	69	0	8	266
>1.10	0	77	0	172	0	77	0	10	336
Totals	8330	8721	10,140	10,851	5863	6787	2048	2656	55,396

Numbers of call and put options T trading days to maturity.

NOTE: The sample period is January 2, 1990, to June 30, 2012 comprising 5672 trading days.

Figure 1. Plot of the S&P 500 Index from January 2, 1990, to June 30, 2012.

Table 9. Parameter estimates for the SV model using the S&P 500 Index from January 2, 1990, to December 30, 2007 and selected options written on the index over that period

Parameter	1–1	0–2	0–4	0–6	0–10
κ	1.8788	1.8805	1.6341	1.5893	1.5429
	(0.3073)	(0.2554)	(0.1880)	(0.1982)	(0.1335)
γ	0.0371	0.0369	0.0471	0.0544	0.0567
	(0.0088)	(0.0047)	(0.0062)	(0.0053)	(0.0037)
σ	0.3857	0.4887	0.6361	0.7072	0.7364
	(0.0882)	(0.0829)	(0.0677)	(0.0502)	(0.0387)
ρ	−0.7410	−0.7648	−0.7510	−0.7313	−0.7225
	(0.1171)	(0.1209)	(0.0709)	(0.0350)	(0.0285)
ξ_s	4.0906	4.3743	4.2748	4.8506	6.4434
	(0.6373)	(0.5182)	(0.7669)	(0.3791)	(0.1837)
ξ_v	−1.9894	−2.1042	−2.1959	−2.0481	−1.8054
	(0.2976)	(0.2302)	(0.2624)	(0.1468)	(0.0976)
α	0.1015	0.0968	0.1090	0.1131	0.1404
	(0.0055)	(0.0048)	(0.0085)	(0.0053)	(0.0052)

NOTE: The estimation procedure uses 35,840 particles and error specification ES4.

Table 10. Parameter estimates for the SV model calculated using the S&P 500 index from January 2, 1990, to December 30, 2011 and selected options written on the index over that period

Parameter	1–1	0–2	0–4	0–6	0–10
κ	1.8729	1.9775	1.8504	1.6655	1.5038
	(0.2260)	(0.2027)	(0.1530)	(0.1529)	(0.1713)
γ	0.0365	0.0376	0.0423	0.0551	0.0594
	(0.0049)	(0.0024)	(0.0030)	(0.0028)	(0.0044)
σ	0.3690	0.4568	0.5923	0.6712	0.7357
	(0.0495)	(0.0488)	(0.0271)	(0.0501)	(0.0308)
ρ	−0.7921	−0.7591	−0.7912	−0.7516	−0.7251
	(0.1126)	(0.0906)	(0.0362)	(0.0425)	(0.0257)
ξ_s	4.2069	3.9403	4.0237	5.2163	4.6670
	(0.5164)	(0.5998)	(0.3948)	(0.3369)	(0.5476)
ξ_v	−2.0523	−1.9629	−2.2906	−1.5935	−1.8409
	(0.2173)	(0.1612)	(0.1954)	(0.1568)	(0.0279)
α	0.0968	0.0946	0.1031	0.1109	0.1321
	(0.0048)	(0.0048)	(0.0051)	(0.0054)	(0.0050)

NOTE: The estimation procedure uses 35,840 particles and error specification ES4.

As in the simulation experiments, the simplex algorithm is used to obtain the maximum likelihood parameter estimates. The values of the parameters used to construct the vertices of the starting simplex are informed by estimates reported in the literature (see, e.g., Tables 11 and 14). When the simplex algorithm converged it was rerun with starting parameter guesses based on results returned by the converged algorithm. This iterative procedure is repeated until two successive iterations returned a similar set of parameter values, as gauged by a user-supplied tolerance. The procedure of restarting the converged estimation from a fully sized simplex affords some additional protection against converging to local optima.

Table 9 reports the results when the distribution of pricing errors is specified in terms of the ES4 assumption in Equation (15). At first the sample is restricted to the period January 2, 1990 to December 30, 2007 to exclude data pertaining to the global financial crisis. This is a relatively stable period over which to check the efficacy of the estimation procedure. Table 10 reports the results for the unrestricted sample covering the period January 2, 1990 to December 30, 2011.

It is immediately apparent that all of the estimated parameters have the expected algebraic sign and all are statistically significant. It is also important to note that the size of the standard errors reported (computed using the outer product of the gradient approximation to the covariance matrix at the optimum) is broadly comparable to the root mean square errors reported in Section 5. This goes to the heart of the debate on the smoothness of the likelihood function obtained from the particle filter. While direct computation of the Hessian is problematic, the function appears smooth enough to allow sensible computation of first derivatives with respect to the parameters. The function accuracy returned here is approximately six decimal places and the gradient is computed by finite differences using a relative error of 10^{-2}. This compares with the default relative error of approximately 10^{-5} when the function value is assumed to be computed to machine precision.

There are perhaps three comments of a general nature to make before turning to a more in-depth comparison of these estimates with estimates published in the existing literature. First, the scaling factors associated with equity and volatility risk premia, namely, ξ_s and ξ_v, are identified and statistically well determined. Second, as generally accepted, the estimate of ρ is negative and significant. Third, the pricing error α grows as more options are added. This result is to be expected, given the consensus view that price profiles (or equivalently implied volatility profiles) observed in the market deviate markedly from those produced by a fixed-parameter SV model. Indeed, daily recalibration of an SV model to observed options data reveals significant time variation in the fitted parameters. As the inclusion of additional options here tends to involve further-from-the-money options, the fit of the SV model to the option profiles deteriorates, placing pressure on the estimated level of price noise.

To check the stability of the parameter estimates of the SV model, estimation was repeated and the period of the global financial crisis included, with the results reported in Table 10. The pattern of the estimates is remarkably similar given the immense turbulence in the financial markets in the years 2008 and 2009. The estimates of the equity and volatility premia, ξ_s and ξ_v, and the correlation parameter, ρ, are however broadly similar to the results reported in Tables 9 and 10. In part, the decision to estimate over two different sub samples is motivated by the study of Jones (2003), who found differences in the risk-neutral dynamics pre- and post-1987. Although no significant differences are apparent here, the issue of dynamics will be returned to presently.

Perhaps the most difficult task in interpreting these results obtained from the particle filter is placing them in the context of previously published results. The difficulty stems from the fact that the published work exhibits differences in notation, model specification, data (including both time-series and cross-sectional differences), and conventions with respect to reporting the parameters. The parameters values reported in this article are annualized values, as are those reported by Aït-Sahalia and Kimmel (2007), Santa-Clara and Yan (2010), and Christoffersen, Jacobs, and Mimouni (2010), but this convention is not uniformly

Table 11. Comparison of parameter estimates obtained from the particle filter implemented in the current article (SV) with previously published values from similar research

Paper	Parameters						
	κ	ρ	σ	γ	$\widetilde{\gamma}$	ξ_s	ξ_v
SV_{2007}	1.634	−0.731	0.636	0.047	exp	4.275	−2.196
SV_{2011}	1.850	−0.752	0.592	0.042	exp	4.024	−2.291
ABL	3.931	−0.597	0.197	0.013		2.560	
CGGT	3.386	−0.279	0.246	0.015			
EJP	5.821	−0.397	0.361	0.023			
ASK	5.130	−0.754	0.520	0.044		2.562	
SCY	18.40	−0.731	0.612	0.046			
CJM	6.520	−0.771	0.460	0.035			
CG	0.931	−0.018	0.062	0.015	0.010		−0.240
Pan	7.100	−0.530	0.320	0.014	exp	8.600	−7.600
Jones	3.704	−0.603	0.524	0.026	0.008		8.216
Jones†	4.561	−0.679	0.227	0.013	exp		−7.408
Eraker	4.788	−0.569	0.554	0.049	0.103		−2.520
CJM‡	2.879	−0.704	0.537	0.063	0.063		0.000

NOTES: Entries marked "exp" represent models with explosive risk-neutral dynamics. The block shaded gray contains studies that do not report estimates for the volatility premium and the final block contains studies that do report estimates for this parameter.
ABL Andersen, Benzoni, and Lund (Andersen et al. (2002)) S&P 500 (1953–1996).
CGGT Chernov et al. (Chernov et al. (2003)) DJIA (1953–1999).
EJP Eraker, Johannes, and Polson (Eraker et al. (2003)) S&P 500 (1980–1999) and NASDAQ100 (1985–1999).
ASK Aït-Sahalia and Kimmel (Aït-Sahalia and Kimmel (2007)) S&P 500 (1990–2004), VIX index.
SCY Santa-Clara and Yan (Santa-Clara and Yan (2010)) S&P 500 (1996–2002), SPX options.
CJM Christoffersen, Jacobs, and Mimouni (Christoffersen et al. (2010)) S&P 500 (1996–2004).
CG Chernov and Ghysels (Chernov and Ghysels (2000)) S&P 500 (1985–1994), SPX options.
Pan Pan (Pan (2002)) S&P 500 (1989–1996), near-the-money short dated options.
Jones Jones (Jones (2003)) S&P100 (1986–2000) and †(1988-2000), VXO index.
Eraker Eraker (Eraker (2004)) S&P 500 (1987–1990), SPX options.
CJM‡ Christoffersen, Jacobs, and Mimouni (Christoffersen et al. (2010)) S&P 500 (1996–2004), SPX call options.

adopted. Table 11 reports the results of previous articles in this field based on the SV model where every attempt has been made to scale, transform, or compute similar quantities to annualized values. This table, therefore, represents an attempt to compare our results with the existing literature within a common framework.

In Table 11, SV is the benchmark square-root stochastic volatility estimated in this article and the 2007 and 2011 are the end dates of the sample used in the estimation. Estimates from previously published work are usually referred to by the authors' initials and a key is provided in the notes to the table. The estimates have been separated into studies, which do not report a volatility premium (studies based primarily on returns data) and those that do report a volatility premium (studies that use returns and options data). The former category is shaded gray in Table 11. Details of the computations involved in producing this comparison are available on request.

An interesting conjecture springs to mind when examining the chronological order of the results in Table 11 and the size of the estimate of the parameter ρ. It is clear that earlier studies find much smaller values of this parameter than that found by more recent studies. This may be attributable either to the characteristics of the time-series data or to the estimation methodology. On balance the range $[-0.75, -0.70]$ reported in this article seems to be consistent with the later literature. The estimated value of ρ will be returned to in the discussion of the results in Table 15 in Section 8. Similarly the estimates of σ and γ are consistent with previous work.

The estimate of κ reported in this study is lower than that of previous studies. However, there is significant variation in the reported values for κ from 0.931 to 18.04 indicating that κ is difficult to resolve. In favor of the estimate presented in this article is that it is generated using a longer time-series of the index (22 years) than many of the previous studies. The articles with comparable span of time-series data are Andersen, Benzoni, and Lund (2002) and Jones (2003) both of whom report values of κ on the low side and which are, broadly speaking, comparable to our estimate.

A feature of the specification of the model and also of the estimation method is that estimates can be provided for both ξ_s and ξ_v. The only other article to provide estimates for both these parameters is Pan (2002). The current results are similar to those of Pan (2002) and Jones (2003) in the respect that the risk-neutral dynamics of the simple SV model are explosive, a situation in which there is no mean reversion to a well-defined long-run volatility. There is little to be gained therefore by exploring this issue any further in the context of this model. It is to be expected that the inclusion of jumps and the associated jump premium may alleviate pressure on the risk-neutral volatility dynamics.

8. ESTIMATING THE SVJ MODEL

Tables 12 and 13 report the results for the estimation of the SVJ model. As in the estimation of the SV model, the estimation is over two sample periods, 1990–2007 and 1990–2011. To

Table 12. Parameter estimates for the SVJ model of stochastic volatility for the S&P 500 Index and selected options written on the index from January 2, 1990, to December 30, 2007

Parameter	1–1	0–2	0–4	0–6	0–10
κ	1.7106	1.6888	1.7173	1.7690	1.5849
	(0.2624)	(0.2086)	(0.2131)	(0.1669)	(0.2308)
γ	0.0492	0.0485	0.0510	0.0502	0.0627
	(0.0153)	(0.0119)	(0.0064)	(0.0038)	(0.0081)
σ	0.6532	0.6586	0.6742	0.6847	0.7522
	(0.1374)	(0.1543)	(0.0609)	(0.0750)	(0.0518)
λ	2.3321	2.2630	2.2573	2.2811	2.2144
	(0.7771)	(0.4755)	(0.2671)	(0.2694)	(0.3668)
$\widetilde{\lambda}$	2.8311	2.7806	2.7535	2.7787	2.7045
	(0.9420)	(0.5074)	(0.4354)	(0.3487)	(0.2952)
δ_J	−0.0205	−0.0199	−0.0194	−0.0200	−0.0054
	(0.0059)	(0.0035)	(0.0028)	(0.0017)	(0.0010)
σ_J	0.0192	0.0214	0.0182	0.0170	0.0166
	(0.0052)	(0.0042)	(0.0018)	(0.0015)	(0.0022)
ρ	−0.7401	−0.7485	−0.7615	−0.7460	−0.7557
	(0.1884)	(0.1341)	(0.0737)	(0.0462)	(0.0311)
ξ_s	2.5918	2.7476	2.7613	2.6677	3.0040
	(0.9591)	(0.5964)	(0.3543)	(0.1836)	(0.4003)
ξ_v	−1.2952	−1.2927	−1.3144	−1.3253	−1.0725
	(0.3572)	(0.2131)	(0.0904)	(0.1186)	(0.1178)
α	0.0961	0.0899	0.1061	0.1195	0.1411
	(0.0209)	(0.0152)	0.0084	(0.0077)	(0.0063)

NOTE: The estimation procedure uses 71,680 particles and error specification ES4.

ensure comparability, exactly the same options are used in the SVJ estimation as those used in the estimation of the SV model.

The pricing errors as measured by the parameter α are slightly lower in the SVJ model. These results suggests that allowing for jump risk is important in obtaining a higher-quality fit to option price profiles. The estimated jump intensities λ and $\widetilde{\lambda}$ indicate between two and three jumps per year, which is consistent with the claim of Broadie, Chernov, and Johannes (2007) that jumps are relatively rare and amount to only one or two a year. It is important to note that the estimates of the risk-neutral intensity are uniformly greater than their physical counterparts, and it is easily verified that the estimates of δ_J and σ_J all imply negative expected values for return jumps, $\mathbb{E}[J-1]$. This, in turn, means that all estimates of the jump premium are found to be positive, given the specification of the jump premium in (8).

As in the case of the SV model, the estimates of ξ_s and ξ_v in the SVJ model are robustly determined being both statistically significant and relatively stable across the choice of number of options. Their values are, however, significantly lower in absolute value than those obtained from the SV model, a result that is consistent with the addition of extra risk factor associated with jumps. Moreover, the model is no longer explosive under the risk-neutral measure, lending added credence to the claim that the SVJ model is a better fit to the data.

Table 14 reports the results of previous articles in this field based on the SVJ model in which a concerted attempt has once again been made to interpret the published results in a common framework. As in Table 11, studies that do not report volatility and jump premia are grouped together and shaded gray to distinguish them from those studies that do report estimates for these parameters. Not much can be added to the discussion of the parameters κ, ρ, σ, and γ, which are not substantively different to the SV model. Broadly speaking, the parameter estimates of the SVJ model are in line with previously published estimates.

One interesting feature of the results reported in Tables 11 and 14 is that the estimates of ρ are remarkably similar. It is tempting to conclude, therefore, that estimation of ρ is relatively robust to the model misspecification of ignoring jumps in returns. A parameter of interest that does appear to be sensitive to this model misspecification, however, is the equity premium parameter, ξ_s. The estimate of ξ_s falls from 4.024 in the SV model to 2.782 in the SVJ model, while the estimate of ξ_s, reported by Pan (2002), falls from 8.6 to 3.6. The estimate reported by Andersen, Benzoni, and Lund (2002) is also relatively unstable, but actually increases from 2.56 to 4.02 when moving from the SV model to the SVJ model.

Where the SVJ model differs from the SV model is in the risk-neutral dynamics and the parameters relating specifically to jumps in equity prices. Consistent with the earlier observation that the current model estimates imply mean-reverting risk-neutral dynamics, this is also a feature of the other two articles that report ξ_v, namely, Pan (2002) and Eraker (2004). The estimate of the intensity parameter, λ, lies in the middle of the range of reported values. To be fair to Pan (2002), however, the jump intensity parameter, λ, reported in Table 14 is distorted because it is scaled by the variance of returns, and it is therefore difficult to draw an immediate comparison between her value of λ and the other reported values in the table.

One aspect of Table 14, which needs some explanation, is the pair of parameters tabulated as μ and ξ_J. It is almost impossible

Table 13. Parameter estimates for the SVJ model of stochastic volatility for the S&P 500 Index and selected options written on the index from January 2, 1990, to December 30, 2011

Parameter	1–1	0–2	0–4	0–6	0–10
κ	1.7603	1.7043	1.7335	1.5735	1.6563
	(0.3301)	(0.2750)	(0.2046)	(0.1926)	(0.1704)
γ	0.0506	0.0515	0.0510	0.0584	0.0588
	(0.0093)	(0.0090)	(0.0059)	(0.0037)	(0.0063)
σ	0.6483	0.6663	0.6630	0.7041	0.7576
	(0.1439)	(0.1021)	(0.0868)	(0.0634)	(0.0464)
λ	2.2576	2.1854	2.2986	2.0931	2.8078
	(0.4962)	(0.2807)	(0.3195)	(0.3186)	(0.2980)
$\widetilde{\lambda}$	2.7479	2.7018	2.8023	2.5984	3.2368
	(0.5228)	(0.4422)	(0.3270)	(0.2982)	(0.5497)
δ_J	−0.0210	−0.0204	−0.0197	−0.0175	−0.0111
	(0.0046)	(0.0039)	(0.0031)	(0.0013)	(0.0019)
σ_J	0.0188	0.0190	0.0186	0.0123	0.0102
	(0.0044)	(0.0039)	(0.0022)	(0.0017)	(0.0014)
ρ	−0.7594	−0.7474	−0.7542	−0.7348	−0.7324
	(0.1574)	(0.1323)	(0.0616)	(0.0435)	(0.0283)
ξ_s	2.7636	2.7448	2.7816	2.9064	2.2661
	(0.7591)	(0.4666)	(0.4201)	(0.1796)	(0.2374)
ξ_v	−1.3145	−1.2874	−1.3227	−1.3520	−1.4821
	(0.3776)	(0.2443)	(0.1705)	(0.1458)	(0.3086)
α	0.0891	0.0825	0.1042	0.1117	0.1297
	(0.0202)	(0.0098)	0.0084	(0.0069)	(0.0052)

NOTE: The estimation procedure uses 71,680 particles and error specification ES4.

Table 14. Comparison of parameter estimates for the SVJ model obtained from the particle filter implemented in the current article with previously published values from similar research

Paper	Parameters									
	κ	ρ	σ	γ	$\widetilde{\gamma}$	ξ_s	ξ_v	μ (%)	ξ_J (%)	λ
SVJ_{2007}	1.717	−0.762	0.674	0.050	0.207	2.761	−1.314	−1.923	0.978	2.257
SVJ_{2011}	1.734	−0.754	0.663	0.051	0.179	2.782	−1.323	−1.953	0.873	2.299
ABL	3.704	−0.620	0.184	0.013		4.020		0.000		5.040
CGGT	2.790	−0.476	0.207	0.016				−2.952		1.700
EJP	3.226	−0.467	0.240	0.021				−2.485		1.512
SCY	18.44	−0.728	0.630	0.036				−3.145		0.674
CJM	6.589	−0.777	0.450	0.032				−1.002		2.790
Pan	6.400	−0.530	0.300	0.015	0.029	3.600	−3.100	−0.800	3.463	12.30
Eraker	4.788	−0.586	0.512	0.042	0.073		−2.016	−0.086	0.805	0.504
CJM‡	2.638	−0.782	0.448	0.063	0.063		0.000	−3.168	1.465	2.832

NOTES: The block shaded gray contains studies that do not report estimates for the volatility and jump premia and the final block contains studies that do report estimates for these parameters.
ABL Andersen, Benzoni, and Lund (Andersen et al. (2002)) S&P 500 (1953–1996).
CGGT Chernov et al. (Chernov et al. (2003)) DJIA (1953–1999).
EJP Eraker, Johannes, and Polson (Eraker et al. (2003)) S&P 500 (1980–1999) and NASDAQ100 (1985–1999).
SCY Santa-Clara and Yan (Santa-Clara and Yan (2010)) S&P 500 (1996–2002), SPX options.
CJM Christoffersen, Jacobs, and Mimouni (Christoffersen et al. (2010)) S&P 500 (1996–2004).
Pan Pan (Pan (2002)) S&P 500 (1989–1996), near-the-money short dated options.
Eraker Eraker (Eraker (2004)) S&P 500 (1987–1990), SPX options.
CJM‡ Christoffersen, Jacobs, and Mimouni (Christoffersen et al. (2010)) S&P 500 (1996–2004), SPX call options.

to find a common specification of the jump size and timing premia for these models across articles. The common theme, however, is the assumption about the discontinuity in returns, which is specified as $\log J \sim N(\delta_J, \sigma_J^2)$, as described in Section 2.2. Consequently, the comparison is made by computing the expected jump size under the physical measure,

$$\mu = \mathbb{E}[J-1] = \exp\left[\delta_J + \sigma_J^2/2\right] - 1\,,$$

and expressing it as a percentage. This quantity is common to all specifications of the SVJ model. Exactly how the premium ξ_J is computed will vary from specification to specification, with its value in this article computed from Equation (8). The estimate of $\mu = -1.953\%$ is quite similar to those reported by Pan (2002) and Christoffersen, Jacobs, and Mimouni (2010), which take the values -0.8% and -3.168%, respectively. Eraker (2004), on the other hand, reported an estimate of -0.086%, which is much closer to zero. The ξ_J value of 0.873% is similar to those reported by Eraker (2004) and Christoffersen, Jacobs, and Mimouni (2010), at 0.805% and 1.465%, respectively, but slightly lower than Pan (2002) who reported 3.463%. This discrepancy can be explained by the fact that Pan (2002) found a lower diffusive component of the equity premium.

As an additional check on the predictions of the SVJ model, the filtered volatility obtained from the (0–6) model in Table 13 is compared with a commonly used proxy for the volatility of the S&P 500, namely, the Implied Volatility Index (VIX) published by the Chicago Board of Exchange. The VIX is constructed from European put and call option prices such that at any given time it represents the risk-neutral expectation of integrated variance averaged over the next 30 calendar days (or 22 trading days). Evidence of the efficacy of the particle filter to track the path of volatility is given by Figure 2, which shows how well the filtered volatility from the estimated model tracks the VIX index, reinforcing the conclusion that the SVJ model has been well estimated.

The final estimation exercise is designed to demonstrate the ability of the particle filter, when implemented taking advantage of the parallel nature of the problem, to handle very large numbers of options. Accordingly, the options are stratified according to their moneyness, namely, the ratio S/K for call options and K/S for put options. The estimation of the models in Table 15 uses all available options within the stratification band of moneyness, which is chosen (approximately 40,000 in the 2007 sample and 60,000 in the 2011 sample). The central issue investigated is whether the conventional wisdom of using out-of-the-money options instead of at- or in-the-money options is confirmed. Accordingly two different stratifications of moneyness are chosen, namely, options that fall within moneyness [0.94, 0.96] and those that fall within moneyness [1.00, 1.02].

It is apparent from the results reported in Table 15 that the models estimated using out-of-the-money options return parameter estimates that accord more closely with the results reported here and those reported in earlier articles, as summarized in Tables 11 and 14. In particular, when using at-the-money options for the estimation, the estimate of ρ shows significantly less leverage and the risk-neutral dynamics have a tendency to become explosive. These observations concerning the use of at- and in-the-money options in estimation may offer a partial explanation for some of the results reported in the early literature concerning the discussion of the estimate of ρ in Section 7. Certainly the tendency for the risk-neutral dynamics to lose their mean-reverting property if at-the-money options are used is an important observation, which deserves further investigation.

One final point concerns the estimate of the pricing error α. The estimate of α is substantially larger for the models using

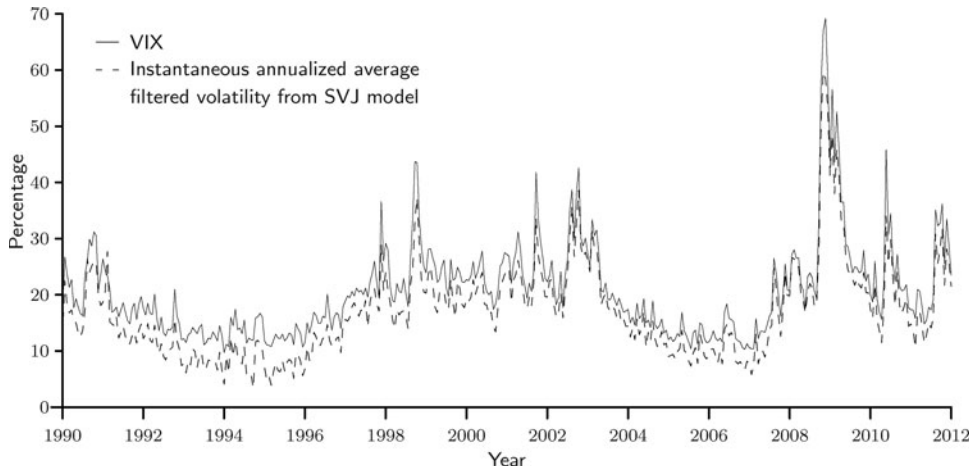

Figure 2. Plot of the VIX versus the instantaneous, annualized filtered volatility obtained from estimates of the SVJ model using the particle filter with 71,680 particles and 6 daily out-of-the-money options.

out-of-the-money options. Recall, however, that the pricing error distribution used here implies that α is a relative error. As noted by Huang and Wu (2004), prices of in-the-money options embed a component related to their intrinsic values that are model independent and thus insensitive to pricing error, it might therefore be expected that the relative pricing errors for out-of-the-money options are of a greater magnitude than their in-the-money counterparts.

Table 15. Parameter estimates for the SVJ model based on S&P 500 with the number of options used in each estimation constrained only by their stratification in terms of spot-to-strike ratio

	1990–2007		1990–2011	
Moneyness	[0.94,0.96]	[1.00,1.02]	[0.94,0.96]	[1.00,1.02]
no. options	41,390	46,394	59,859	58,003
κ	1.5982	1.3878	1.7460	1.2214
	(0.1886)	(0.0960)	(0.3294)	(0.0411)
γ	0.0424	0.0436	0.0443	0.0507
	(0.0045)	(0.0027)	(0.0067)	(0.0026)
σ	0.5574	0.2430	0.7084	0.2640
	(0.0372)	(0.0145)	(0.0535)	(0.0113)
λ	1.6115	2.3907	2.2552	0.3609
	(0.1839)	(0.1181)	(0.5019)	(0.0139)
$\widetilde{\lambda}$	2.0975	2.8293	2.7831	1.2695
	(0.2089)	(0.1590)	(0.3753)	(0.0720)
δ_J	−0.0077	−0.0165	−0.0178	−0.0113
	(0.0009)	(0.0009)	(0.0025)	(0.0005)
σ_J	0.0338	0.0124	0.0253	0.0054
	(0.0018)	(0.0007)	(0.0028)	(0.0002)
ρ	−0.8746	−0.2653	−0.7283	−0.5097
	(0.0362)	(0.0198)	(0.0255)	(0.0207)
ξ_s	2.8319	2.3057	2.9255	2.4783
	(0.2532)	(0.1932)	(0.5334)	(0.0906)
ξ_v	−1.3188	−1.1958	−1.3276	−1.3779
	(0.1282)	(0.0516)	(0.2814)	(0.0591)
α	0.2400	0.0648	0.2346	0.0635
	(0.0193)	(0.0031)	(0.0157)	(0.0021)

NOTE: The estimation uses 71,680 particles in combination with the error specification ES4.

9. CONCLUSION

Observations on option prices contain a wealth of information pertinent to the parameter estimation problem and the associated filtering of the latent state variable(s) particularly because of their dependence upon the parameters of the system under the risk-neutral measure. When used together with index returns data, which depend solely on the model parameters under the physical measure, powerful advances can be made. Although considerable work has been undertaken to incorporate options data into estimation, to date, however, the computational burden imposed by the associated range of estimation procedures has significantly impeded progress. Although there are a few studies in which parameter estimates have been successfully provided by Markov chain Monte Carlo methods, the ability to provide maximum likelihood estimates using a particle filter have to date been thwarted by computational considerations. The main contribution of this article is the use of graphics processor technology and parallel computation to allow simultaneous estimation of all parameters (both physical and risk-neutral) of the model using a lengthy time-series of the underlying stock market index and very large numbers of options written on the index.

A simulation study designed to test the efficacy of the particle filter and explore the performance of alternative assumptions about the distribution of the pricing errors of options is performed. The results confirm that the filter works extremely well, although little concrete guidance is provided as to the appropriate choice of pricing error distribution. When the stochastic volatility models are estimated using S&P 500 index and options data, the parameter estimates returned are consistent with those reported by many previous studies. In sharp contrast to many of these previous studies, however, statistically significant equity and volatility premia are identified. Moreover, in the stochastic volatility model that incorporates jumps in equity prices, the values of the parameters returned by the filter indicate that the dynamics of volatility are not explosive under the risk-neutral measure. We take this to provide evidence of the internal consistency of the model and the estimates. Finally, the power of the method is demonstrated with an estimation problem that uses all the available options written on the S&P 500 index limited only by a requirement to satisfy a certain degree of moneyness. The numbers of options involved in the estimation range from 41,390 to 59,859, which is an indication of the potential offered by this approach.

ACKNOWLEDGMENTS

The authors acknowledge the Multi-modal Australian ScienceS Imaging and Visualisation Environment (MASSIVE) and also Mark Barry and Ashley Wright from the QUT's High Performance Computing Group. The authors thank the associate editor and two anonymous referees for their constructive comments. This research was supported under Australian Research Council's Discovery Projects funding scheme (DP120100837).

[Received July 2013. Revised September 2014.]

REFERENCES

Aït-Sahalia, Y., and Kimmel, R. (2007), "Maximum Likelihood Estimation of Stochastic Volatility Models," *Journal of Financial Economics*, 83, 413–452. [579,581,589]

Andersen, T. G., Benzoni, L., and Lund, J. (2002), "An Empirical Investigation of Continuous-time Models for Equity Returns," *Journal of Finance*, 57, 1239–1284. [590,592]

Andersen, T. G., Fusari, N., and Todorov, V. (2015), "Parametric Inference and Dynamic State Recovery From Option Panels," *Econometrica*, 83, 1081–1145. [579]

Bakshi, G., Cao, C., and Chen, Z. (1997), "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52, 2003–2049. [579]

Bates, D. S. (1996), "Jump and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *Review of Financial Studies*, 9, 69–107. [579,581,582]

——— (2000), "Post-'87 Crash Fears in S&P 500 Futures Options," *Journal of Econometrics*, 94, 181–238. [579]

Brandt, M. W., and Santa-Clara, P. (2002), "Likelihood Estimation of Diffusions With an Application to Exchange Rate Dynamics in Incomplete Markets," *Journal of Financial Economics*, 63, 161–210. [584]

Broadie, M., Chernov, M., and Ghysels, M. (2007), "Model Specification and Risk Premia: Evidence From Futures Options," *Journal of Finance*, 62, 1453–1490. [579,591]

Carr, P., and Madan, D. B. (1999), "Option Valuation Using the Fast Fourier Transform," *Journal of Computational Finance*, 2, 61–73. [582,583]

Chernov, M., Gallant, A. R., Ghysels, E., and Tauchen, G. (2003), "Alternative Models for Stock Price Dynamics," *Journal of Econometrics*, 116, 225–257. [590,592]

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. [579,590]

Christoffersen, P., and Jacobs, K. (2004), "The Importance of the Loss Function in Option Valuation," *Journal of Financial Economics*, 72, 291–318. [582]

Christoffersen, P., Jacobs, K., and Mimouni, K. (2010), "Volatility Dynamics for the S&P500: Evidence From Realized Volatility, Daily Returns and Option Prices," *Review of Financial Studies*, 23, 3141–3189. [579,589,592]

Duffie, D., Pan, J., and Singleton, K. (2000), "Transform Analysis and Asset Pricing for Affine Jump-Diffusions," *Econometrica*, 68, 1343–1376. [582]

Durham, G., and Geweke, J. (2011), "Massively Parallel Sequential Monte Carlo for Bayesian Inference," unpublished manuscript. [580]

——— (2013), "Adaptive Sequential Posterior Simulators for Massively Parallel Computing Environments," Working Paper Series 9, Economics Discipline Group, UTS Business School, University of Technology, Sydney, Australia. [580]

Elerian, O. (1998), "A Note on the Existence of a Closed Form Conditional Transitional Density for the Milstein Scheme," Working Paper 1998–W18, Nuffield College, Oxford. [584]

Eraker, B. (2004), "Do Stock Prices and Volatility Jump? Evidence From Spot and Option Prices," *Journal of Finance*, 59, 1367–1403. [579,585,587,590,592]

Eraker, B., Johannes, M., and Polson, N. (2003), "The Impact of Jumps in Equity Index Volatility and Returns," *Journal of Finance*, 58, 1269–1300. [579,590,592]

Fang, F., and Oosterlee, C. W. (2008), "A Novel Pricing Method for European Options Based on Fourier-Cosine Series Expansions," *SIAM Journal on Scientific Computing*, 31, 826–848. [583]

Florens-Zmirou, D. (1989), "Approximate Discrete-Time Schemes for Statistics of Diffusion Processes," *Statistics*, 20, 547–557. [584]

Forbes, C. S., Martin, G. M., and Wright, J. (2007), "Inference for a Class of Stochastic Volatility Models Using Option and Spot Prices: Application of a Bivariate Kalman Filter," *Econometric Reviews, Special Issue on Bayesian Dynamic Econometrics*, 26, 387–418. [579,582]

Gill, P. E., Murray, W., and Wright, M. H. (1981), Practical Optimization, New York: Academic Press. [584]

Heston, S. L. (1993), "A Closed-Form Solution for Options With Stochastic Volatility With Applications to Bond and Currency Options," *Review of Financial Studies*, 6, 327–343. [579,580,582]

Huang, J.-Z., and Wu, L. (2004), "Specification Analysis of Option Pricing Models Based on Time-Changed Levy Processes," *Journal of Finance*, 59, 1405–1439. [582,587,593]

Hurn, A. S., Lindsay, K. A., and McClelland, A. J. (2013), "Quasi-maximum Likelihood Estimation of the Parameters of Multivariate Diffusions," *Journal of Econometrics* 172, 106–126. [584]

Jensen, B., and Poulsen, R. (2002), "Transition Densities of Diffusion Processes: Numerical Comparison of Approximation Techniques," *Journal of Derivatives*, 9, 18–32. [584]

Johannes, M. S., Polson, N. G., and Stroud, J. R. (2009), "Optimal Filtering of Jump Diffusions: Extracting Latent States From Asset Prices," *Review of Financial Studies*, 22, 2759–2799. [579]

Jones, C. S. (2003), "The Dynamics of Stochastic Volatility: Evidence From Underlying and Options Markets," *Journal of Econometrics*, 116, 181–224. [579,589,590]

Lewis, A. L. (2001), "A Simple Option Formula for General Jump-diffusion and Their Exponential Levy Processes," unpublished manuscript. [583]

Lord, R., Fang, F., Bervoets, F., and Oosterlee, C. W. (2008), "A Fast and Accurate FFT-based Method for Pricing Early-Exercise Options Under Lévy Processes," *SIAM Journal on Scientific Computing*, 30, 1678–1705. [583]

Nelder, J. A., and Mead, R. (1965), "A Simplex Method for Function Minimization," *Computer Journal*, 7, 308–313. [584]

Pan, J. (2002), "The Jump-Risk Premia Implicit in Options: Evidence From an Integrated Time-Series Study," *Journal of Financial Economics*, 63, 3–50. [579,581,587,590,592]

Pedersen, A. R. (1995), "A New Approach to Maximum Likelihood Estimation for Stochastic Differential Equations Based on Discrete Observations," *Scandinavian Journal of Statistics*, 22, 55–71. [584]

Pitt, M. K. (2002), "Smooth Particle Filters for Likelihood Evaluation and Maximisation," Working Paper, Warwick Economic Research Papers, No. 651. [580]

Santa-Clara, Y., and Yan, S. (2010), "Crashes, Volatility, and the Equity Premium: Lessons From S&P500 Options," *Review of Economics and Statistics*, 92, 435–451. [579,589,592]

Shoji, I., and Ozaki, T. (1997), "Comparative Study of Estimation Methods for Continuous Time Stochastic Processes," *Journal of Time Series Analysis*, 18, 485–506. [584]