ESTIMATION METHODS FOR STOCHASTIC VOLATILITY MODELS: A SURVEY

Carmen Broto and Esther Ruiz

Departamento de Estadística, Universidad Carlos III de Madrid

Abstract. Although stochastic volatility (SV) models have an intuitive appeal, their empirical application has been limited mainly due to difficulties involved in their estimation. The main problem is that the likelihood function is hard to evaluate. However, recently, several new estimation methods have been intro duced and the literature on SV models has grown substantially. In this article, we review this literature. We describe the main estimators of the parameters and the underlying volatilities focusing on their advantages and limitations both from the theoretical and empirical point of view. We complete the survey with an application of the most important procedures to the S&P 500 stock price index.

Keywords. Bayesian procedures; GMM; Indirect inference; Kalman filter; Lever age effect; Long memory; Maximum likelihood; Monte Carlo Markov Chain; QML; SV M

1. Introduction

During the last decade, there has been an increasing interest in modelling the dynamic evolution of the volatility of high-frequency series of financial returns using stochastic volatility (SV) models. The main characteristic of these models is that the volatility is modelled as an unobserved latent variable. SV models are attractive because they are close to the models often used in Financial Theory to represent the behaviour of financial prices. Furthermore, their statistical properties are easy to derive using well-known results on log-normal distributions. Finally, compared with the more popular GARCH models, they capture in a more appropriate way the main empirical properties often observed in daily series of financial returns (see, for example, Carnero *et al.*, 2003). However, until recently, their empirical applications have been very limited mainly because the exact like- lihood function is difficult to evaluate and Maximum Likelihood (ML) estimation of the parameters is not straightforward. However, lately, several estimators who overcome this problem have been proposed and the literature on SV models has grown substantially. The objective of this article is to review this literature.

The article is organized as follows. Section 2 describes the properties of the basic Gaussian Autoregressive SV (ARSV) model and several of its more popular

extensions as, for example, leverage effect, fat-tailed errors, long-memory and multivariate models. Section 3 describes the main advantages and limitations of several estimators of the parameters and underlying volatilities. This section is completed with an illustration with simulated data. Section 4 describes several estimation methods for Long-Memory SV (LMSV) models. Section 5 illustrates the results with an empirical application that compares the principal estimators by fitting the ARSV model to daily S&P 500 returns. Finally, section 6 concludes the article.

2. Stochastic Volatility Models

2.1 Autoregressive Stochastic Volatility Model

SV models are closely related to financial models often used to represent stock prices [see, for example, Ghysels et $al.$ (1996) and Barndorff-Nielsen et $al.$ (2002)]. In the context of Financial Econometrics, they were first introduced by Taylor (1986) who proposes to model the logarithm of volatility as an AR(1) process. The corresponding model, known as ARSV(1), is given by

$$y_t = \sigma_* \varepsilon_t \sigma_t \tag{1}$$

$$\log \sigma_t^2 = \phi \log \sigma_{t-1}^2 + \eta_t, t = 1, ..., T \tag{2}$$

where y_t is the return observed at time t and σ_t is the corresponding volatility. σ_* is a scale parameter that removes the necessity of including a constant term in the equation of the log-volatility and ε_t is a white noise process with unit variance that represents the innovations in the level of returns. The disturbance of the volatility equation, η_t, is assumed to be a Gaussian white noise processes with variance σ_η^2, independent of ε_t. Although the assumption of Gaussianity of η_t can seem ad hoc at first sight, Andersen et $al.$ (2001, 2003) show that the log-volatility process can be well approximated by a Normal distribution. The variance of the log-volatility process, σ_η^2, measures the uncertainty about future volatility. If $\sigma_\eta^2 = 0$, the ARSV(1) model is not identified. As we will see later, the parameter ϕ is often considered as a measure of the persistence of shocks to the volatility. Notice that, when ϕ is close to 1 and σ_η^2 is close to 0, the evolution of volatility over time is very smooth. In the limit, if $\phi = 1$ and $\sigma_\eta^2 = 0$, the volatility is constant over time and, consequently, returns are homoscedastic. Harvey and Streibel (1998) propose to test for homoscedasticity using this property of the ARSV model.

The properties of ARSV models have been summarized by Taylor (1994), Shephard (1996), Ghysels et $al.$ (1996), Capobianco (1996) and Barndorff-Nielsen and Shephard (2001). In particular, it is easy to prove that the process y_t is a martingale difference. If $|\phi| < 1$, then σ_t^2 is stationary and also y_t. In this case, the marginal variance of returns is given by

$$\sigma_y^2 = \sigma_*^2 \exp\{0.5\sigma_h^2\} \tag{3}$$

where $\sigma_h^2 = \sigma_\eta^2/(1 - \phi^2)$. is the marginal variance of the log-volatility process. The kurtosis of y_t is given by

$$\kappa_y = \kappa_\varepsilon \exp(\sigma_h^2) \tag{4}$$

where κ_ε is the kurtosis of ε_t. Notice that, if κ_ε is finite, the kurtosis exists if $|\phi| < 1$. Therefore, as far as the model is stationary, the dynamic evolution of the volatility is not further restricted to guarantee the existence of the fourth order moment. Bai et al. (2003) show that the contribution to the overall kurtosis of returns of the kurtosis of the innovations ε_t and the kurtosis due to the volatility clustering is symmetric.

Although the series y_t is uncorrelated, it is not an independent sequence. The dynamics of the series appear in the squared returns. Their autocorrelation function (acf), derived by Taylor (1986), is given by

$$\rho_2(\tau) = \frac{\exp(\sigma_h^2 \rho_h(\tau)) - 1}{\kappa_\varepsilon \exp(\sigma_h^2) - 1}, \quad \tau \geq 1, \tag{5}$$

where $\rho_h(\tau)$ is the autocorrelation of order τ of the underlying log-volatility that, in model (1) and (2), is given by $\rho_h(\tau) = \phi^\tau$. If σ_h^2 is small and/or ϕ close to 1, Taylor (1986) shows that the acf in (5) can be approximated by

$$\rho_2(\tau) \simeq \frac{\exp(\sigma_h^2) - 1}{\kappa_\varepsilon \exp(\sigma_h^2) - 1} \phi^\tau \tag{6}$$

The pattern of the approximated autocorrelations in (6) is the same as for the autocorrelations of an ARMA(1,1) model. However, Carnero et al. (2003) show that the behaviour of the autocorrelations in (5) and (6) can be rather different. In any case, the parameter ϕ has been considered as the measure of the persistence of the autocorrelations of squared returns.

Although it is rather common to assume that the errors ε_t have a Gaussian distribution, several authors have also considered heavy-tailed distributions [see, for example, Nelson (1988), Harvey et al. (1994), Barndorff-Nielsen (1997), Gallant et al. (1997), Mahieu and Schotman (1998), Sandmann and Koopman (1998), Steel (1998), Liesenfeld and Jung (2000), Anderson (2001) and Watanabe and Asai (2001) among many others]. Notice that even when ε_t is assumed to be Gaussian, the distribution of y_t conditional on past observations up to time $t-1$ is not Gaussian.

It is also interesting to mention that Liesenfeld and Richard (2003) show that the ARSV model with Student-t errors can be interpreted as a SV model with two independent volatility processes [see also Shephard (1996) and Chernov et al. (2003) for SV models where the volatility has different components].

The ARSV(1) model can be linearized by taking logarithms of squared returns:

$$\log(y_t^2) = \mu + h_t + \xi_t \tag{7}$$

$$h_t = \phi h_{t-1} + \eta_t \tag{8}$$

where $\mu = \log(\sigma_*^2) + E(\log(\varepsilon_t^2))$, $h_t = \log(\sigma_t^2)$ and $\xi_t = \log(\varepsilon_t^2) - E(\log(\varepsilon_t^2))$. Model (7) and (8) is a non-Gaussian linear state space model where (7) is the measurement equation and (8) is the transition equation. The properties of the measurement noise, ξ_t, depend on the distribution of ε_t. If, for example, ε_t is Gaussian, ξ_t has a $\log \chi_{(1)}^2$ distribution with $\sigma_\xi^2 = Var(\log(\varepsilon_t^2)) = \frac{\pi^2}{2}$ and $E(\xi_t^4) = 3\sigma_\xi^2 + \pi^4$ [see Abramowitz and Stegun (1970)].

Given that $\log(y_t^2)$ is the sum of two independent stationary processes, it is easy to derive its acf:

$$\rho(\tau) = \frac{\sigma_h^2}{\sigma_h^2 + \sigma_\xi^2} \phi^\tau \tag{9}$$

Pérez and Ruiz (2003) have analysed, in the context of SV models, the properties of the sample autocorrelations of y_t^2 and $\log(y_t^2)$ as estimators of the corresponding population counterparts.

Recently Yu *et al.* (2002) propose a new class of SV models, the non-linear SV model, where the volatility is transformed according to the Box–Cox power function. This new specification includes the log-normal SV model.

An interesting extension of the basic ARSV(1) model consists of incorporating the asymmetric response of volatility to negative and positive returns. This effect, known as leverage effect, was first pointed out by Black (1986) and has been often observed in real time series of returns. Harvey and Shephard (1996) proposed to introduce the leverage effect into the ARSV model by allowing the disturbances ε_t and η_{t+1} to be negatively correlated. Later, Jacquier *et al.* (2004) propose an ARSV model with leverage effect where ε_t and η_t are correlated. However, Yu (2002b) points out the problems involved in the latter specification and provides empirical evidence favouring the specification proposed by Harvey and Shephard (1996). Finally, So *et al.* (2002) propose a threshold SV model that is also able to represent the leverage effect.

On the other hand, it could be expected that higher levels of volatility are associated with higher expected returns. The SV in mean (SVM) model incorporates this relationship between mean returns and volatilities. The SVM model as proposed, for example, by Koopman and Uspensky (2002) is given by

$$y_t = \delta\sigma^{*2}\sigma_t^2 + \sigma^*\varepsilon_t\sigma_t \tag{10}$$

where σ_t is defined as in equation (2) and δ is the parameter measuring the volatility-in-mean effect.

Finally, it has often been observed that the decay towards zero of the acf of the squared returns is neither exponential, as in short-memory processes, nor implies

a unit root, as in integrated processes. Consequently, it has been suggested that the volatility may be modelled as a long memory process, whose auto-correlations decay at an hyperbolic rate [see, for example, Ding *et al.* (1993)]. In the context of SV models, Breidt *et al.* (1998) and Harvey (1998) propose independently the LMSV model, where the log-volatility is modelled as an ARFIMA (p,d,q) process. In particular, if $p=1$ and $q=0$, the volatility is given by

$$(1 - \phi L)(1 - L)^d \log(\sigma_t^2) = \eta_t \tag{11}$$

where d is the long memory parameter. Model (11) is stationary when $|\phi| < 1$ and $d < 0.5$ and encompasses the short memory model in (2) when $d=0$. Harvey (1998) shows that the acf of squared returns is given by expression (5) where the variance and the acf of the underlying log-volatility are given by

$$\sigma_h^2 = \sigma_\eta^2 \frac{\Gamma(1-2d)}{[\Gamma(1-d)]^2} \frac{F(1, 1+d; 1-d; \phi)}{(1+\phi)}$$

$$\rho_h(\tau) = \left(\prod_{i=0}^{\tau-1} \frac{d+i}{1-d+i}\right) \frac{F(1, d+\tau; 1-d+\tau; \phi) + F(1, d-\tau; 1-d-\tau; \phi) - 1}{(1-\phi)F(1, 1+d; 1-d; \phi)}$$

respectively, where $\Gamma(\cdot)$ is the gamma distribution and $F(\cdot, \cdot; \cdot; \cdot)$ is the hyper-geometric function. Andersen and Bollerslev (1997) and Robinson (2001) show that, for large lags, τ, the autocovariances of squares decay at the same rate as the autocovariances of h_t. However, if τ is small, the rate of decay of the acf in (5) could be rather different from the rate of decay of the acf of h_t (Pérez & Ruiz, 2003).

2.2 *Multivariate Stochastic Volatility Models*

Multivariate models have been often used to represent systems of financial returns related, for example, with the Asset Pricing Theory (APT), asset allocation, estimation of time-varying betas or Value-at-Risk (VaR). In the context of stochastic volatility models, Harvey *et al.* (1994) propose the following multivariate model which allow the variances and covariances to evolve through time with possibly common trends

$$y_{it} = \sigma_{*i} \varepsilon_{it} \sigma_{it}, i = 1, \ldots, n \tag{12}$$

$$\log(\sigma_{it}^2) = \phi_i \log(\sigma_{it\ 1}^2) + \eta_{it}$$

where $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})'$ has a multivariate Normal distribution with covariance matrix Σ_ε and $\eta_t = (\eta_{1t}, \ldots, \eta_{nt})'$ is distributed independently from ε_t following a multivariate Normal distribution with covariance matrix Σ_η.

The main limitation of this model is that it restricts the correlations to be constant over time. Danielsson (1998) extends this model to allow for leverage

effects and time-varying correlations. Later, Ray and Tsay (2000) used the same model to study common long memory components in daily stock volatilities of groups of companies.

Alternatively, Jacquier *et al.* (1995) and Shephard (1996) propose a factor model for returns, where the factors are SV processes. The factor-SV model is given by

$$Y_t = Df_t + a_t, \quad t = 1, \ldots T \tag{13}$$

$$f_{it} = \varepsilon_{it}\sigma_{it}, \quad i = 1, \ldots, K$$

$$\log(\sigma_{it}^2) = \omega_i + \phi_i \log(\sigma_{it\ 1}^2) + \eta_{it}$$

where $Y_t = (y_{1t}, \ldots, y_{nt})$ is the vector of returns observed at time t, D is the matrix of factor loadings restricted as usual to be identified and $a_t = (a_{1t}, \ldots, a_{nt})$ is the vector of idiosyncratic noises assumed to be NID(0, Σ_a) where $\Sigma_a = diag(\sigma_{ai}^2)$. Jacquier *et al.* (1995), Shephard (1996) and Liesenfeld and Richard (2003) have considered a single-factor model, i.e. $K=1$. Extensions to K factors and/or idiosyncratic errors following SV processes have been analysed by Kim *et al.* (1998), Pitt and Shephard (1999b) and Aguilar and West (2000). Finally, Chib *et al.* (2002b) generalize model (13) allowing series-specific jumps at each time and Student-t innovations, a_t, with unknown degrees of freedom, and Lopes and Migon (2003) allow for factor loadings to evolve over time.

Very recently, Tims and Mahieu (2003) have proposed to fit a multivariate SV model to the logarithmic range. The logarithmic range is a proxy for volatility that is approximately Gaussian, allowing for the application of standard Kalman filter techniques to estimate the corresponding models.

3. Estimation of the Parameters of ARSV(1) Models

Estimation methods of the parameters of the short-memory ARSV(1) model in (1) and (2) can be classified into two general groups. The first class is comprised by estimators based directly on the statistical properties of y_t. There are three main classes of estimators within this group: (i) estimators based on the method of moments (MM); (ii) estimators based on the ML principle and (iii) estimators based on an auxiliary model. Alternatively, the parameters of the ARSV(1) model can also be estimated using the linear model in (7) and (8). In subsection 3.1, we describe the estimators within the first group while subsection 3.2 deals with the estimators based on $\log(y_t^2)$.

3.1 *Methods Based on y_t*

3.1.1 *Method of Moments*

The simplest estimator within this group is the MM used by, for example, Taylor (1986). Later, Melino and Turnbull (1990) proposed to estimate the parameters of the ARSV(1) model using generalized method of moments (GMMs) which, under

very general conditions, is consistent and asymptotically normal (Hansen, 1982). These estimators are based on the convergence of selected sample moments to their unconditional expected values. The procedure also implies the estimation of a weighting matrix that takes into account the non-iid property of the moment discrepancies. Andersen and Sørensen (1996) propose to improve the MM procedure by using a penalty function to ensure stationarity of the model and employing a modified weighting matrix. The problem is that the score function in ARSV models cannot be computed and, consequently, the proper set of moments to be used should be guessed. Alternatively, Duffie and Singleton (1993) proposed the simulated method of moments (SMM) that replaces the analytical moments by the moments of a simulated process.

The main attraction of these estimators is that their empirical implementation is very simple. Consequently, they have been extensively used to estimate the parameters of SV models [see, for example, Andersen (1994b), Ghysels and Jasiak (1996), Vetzal (1997) and Fleming *et al.*, 1998].

However, these procedures have poor finite sample properties, and their efficiency is suboptimal with respect to ML methods. Jacquier *et al.* (1994) (JPR) find substantial bias and sampling variability, especially for the estimates of σ_η^2 and show that the performance of the technique worsens when ϕ is close to 1 and the coefficient of variation, defined as $CV = V(\sigma_t^2)/(E(\sigma_t^2))^2 = \exp\left\{\frac{\sigma_\eta^2}{1-\phi^2}\right\} - 1$, is small. As real time series of returns are characterized by highly persistent volatilities, it seems that the methods based on the MM principle are not suitable for estimating the parameters of SV models. Furthermore, the GMM criterion surface for the ARSV(1) model is highly irregular. Therefore, optimization fails to converge, specially for small sample sizes. A large amount of non-converging estimations has been reported by Andersen and Sørensen (1996). This problem can be caused by imprecise estimates of the long-run covariance matrix and, consequently, would disappear with an accurate approximation of the true weighting matrix (Andersen *et al.*, 1999).

Finally, given that GMM estimators do not generate estimates of the underlying volatilities, other procedures as, for example, the Kalman filter or the reprojection technique of Gallant and Tauchen (1998) should be used to get them [see, for example, Andersen (1994b) and Ghysels and Jasiak (1996)].

3.1.2 *Maximum Likelihood Estimators*

Recently, ML estimators of the parameters of SV models have experienced a big progress, thanks to the development of numerical methods based on importance sampling and Monte Carlo Markov Chain (MCMC) procedures. In order to derive the likelihood, the vector of the unobserved volatilities has to be integrated out of the joint probability distribution. If we denote by $Y_T = (y_1, \ldots, y_T)'$ the vector of observations, $\sigma = (\sigma_1, \ldots, \sigma_T)'$ the vector of unobserved volatilities and $\theta = (\omega, \phi, \sigma_\eta^2)'$ the vector of unknown parameters, the likelihood is given by

$$f(Y_T|\theta) = \int f(Y_T|\sigma, \theta)f(\sigma|\theta)d\sigma \qquad (14)$$

The dimension of the integral in (14) is equal to the sample size, T, and its evaluation requires numerical methods.

The first importance sampling algorithm applied to SV models was by Geweke (1994), who also introduced the possibility of using MCMC in this setting. Although his procedure has not been directly applied to real time series, it is important because it was the basis of many subsequent procedures, for example, Danielsson (1994), Shephard and Pitt (1997) and Sandmann and Koopman (1998). The main attraction of importance sampling over MCMC algorithms is that it is less computationally demanding and avoids convergence problems. Furthermore, the accuracy of the estimators can be increased by augmenting the simulation sample size. On the other hand, MCMC algorithms are more flexible because they allow large dimensional problems to be split into smaller dimensional tasks.

MCMC estimators of the parameters of ARSV(1) models, proposed independently by Shephard (1993) and Jacquier *et al.* (1994), have been extensively used in the context of SV models [see, for example, So *et al.* (1998), Steel (1998), Mahieu and Bauer (1998), Chib *et al.* (2002a), So *et al.* (2002) and Yu *et al.* (2002) among many others]. The Bayesian approach for estimating the parameters θ is to augment this vector with the latent log-volatilities, $H_T = (\log(\sigma_1^2), \ldots, \log(\sigma_T^2))'$, forming the vector $\delta = (\theta', H_T')' = (\delta_1, \delta_2, \ldots, \delta_m)'$ Then, the problem is to obtain draws from the posterior density $f(\delta|Y_T)$. The MCMC algorithm creates a Markov process on the blocks of the unknown parameters and latent volatilities and run the simulation sufficiently long, so that the distribution of the current values of the process converges to the posterior density. There are several Markov chain methods that can lead to the same posterior distribution, and the Gibbs sampling is perhaps the most popular. Given a starting value $\delta^{(0)}$, the algorithm consists on obtaining draws $\delta_1^{(i)}$ from $f(\delta_1|\delta_2^{(i-1)}, \ldots, \delta_m^{(i-1)}, Y_t)$, $\delta_2^{(i)}$ from $f(\delta_2|\delta_1^{(i)}, \delta_3^{(i-1)} \ldots, \delta_m^{(i-1)}, Y_t)$, $\ldots$, and $\delta_m^{(i)}$ from $f(\delta_m|\delta_1^{(i)}, \delta_2^{(i)} \ldots, \delta_{m-1}^{(i)}, Y_t)$ from $i = 1, \ldots, M + N$. Tierney (1994) shows that under mild conditions, $(\delta_1^{(i)}, \ldots, \delta_m^{(i)})$ converges in distribution to the posterior as i tends to ∞. Usually, the first M burn-in iterates are discarded and the last N iterates are considered as an approximately independent sample from $f(\delta|Y_T)$. Therefore, the MCMC procedure requires to simulate from suitable conditional distributions. In these iterations, sampling from $f(\theta|H_T, Y_T)$ is fairly straightforward, if conjugate priors are used. On the other hand, sampling from $f(H_T|Y_T, \theta)$ can be done by a single-state sampler where each observation h_t is drawn conditional on the value of all the other observations. However, if ϕ is close to 1 and σ_η^2 is small, the volatility is highly correlated and the single-state sampler, as used by Jacquier *et al.* (1994), generates high level of correlation between draws and, consequently, the algorithm is inefficient and converges very slowly to the posterior distribution (Shephard & Kim, 1994). The efficiency of a MCMC algorithm is determined by how close these

potentially correlated draws from the posterior distribution are from being independent. Sampling the latent volatility can be done more efficiently with multi-state samplers that draw the entire vector H_T at once. Shephard and Pitt (1997) have suggested an efficient multi-move algorithm that outperforms the single block approach in computational terms (Sandmann & Koopman, 1998; Wong, 2002). Multi-move samplers have also been considered by Kim *et al.* (1998), So *et al.* (1998), So and Li (1999), Chib *et al.* (2002a) and So *et al.* (2002).

In any case, after M Monte Carlo replicates of the parameters have been obtained, $\theta^{(i)}$, it is possible to obtain density estimates. Point estimates of any function $g(\theta)$ can also be formed using the posterior mean, i.e. $\widehat{g(\theta)} = \frac{1}{N}\sum_{i=M+1}^{M+N} g(\theta^{(i)})$. On the other hand, a natural choice to obtain smooth estimates of the log-volatilities is the marginal posterior expectation which can be estimated by the sample mean given by $\frac{1}{N}\sum_{i=M+1}^{M+N} h_t^{(i)}$. Finally, it is also possible to obtain interval predictions of future volatilities conditional on the information available at time T that take into account the inherent model variability and the parameter uncertainty.

In the Bayesian framework, comparison between alternative models can be carried out using the Bayes factor which involves the calculation of the marginal likelihood [see, for example, Kim *et al.* (1998), Jacquier *et al.* (2004) or Yu (2002b) for alternative procedures to compute the marginal likelihood]. Berg *et al.* (2004) propose to compare models via the information criteria.

Summarizing, the main attraction of MCMC procedures is that they permit to obtain simultaneously sample inference about the parameters, smooth estimates of the unobserved variances and predictive distributions of multistep forecasts of volatility. On top of that, if the simulation size is large, they have asymptotically the same distribution as the ML estimator. In any case, notice that an important advantage of the MCMC estimators is that inference is based on finite sample distributions and, consequently, the asymptotic approximation is not needed. The main disadvantage is that their implementation is rather complicate and computer consuming.

By means of an extensive Monte Carlo study, Jacquier *et al.* (1994) show that MCMC is more efficient than both quasi-maximum likelihood (QML)[1] and GMM estimators. They consider the following parameter design that has been extensively used afterwards in the literature: $\phi = 0.9$ with $\sigma_\eta^2 = (0.4556, 0.1317, 0.0182)$; $\phi = 0.95$ with $\sigma_\eta^2 = (0.2337, 0.0676, 0.0092)$ and, finally, $\phi = 0.98$ with $\sigma_\eta^2 = (0.0948, 0.0274, 0.0037)$.

Recently, MCMC procedures have been extended to estimate the parameters and volatilities of SV models with fat tails [see, for example, Watanabe and Asai (2001), Chib *et al.* (2002a), Jacquier *et al.* (2004) and So *et al.* (2002)]. SV models with leverage effect have been estimated using MCMC procedures by Meyer and Yu (2000), Yu (2002b), So *et al.* (2002) and Eraker *et al.* (2003). However, Andersen (1994a) points out that Jacquier *et al.*'s (1994) procedure is not easily applicable to SVM models. In general, due to its lack of flexibility, these procedures need to be non-trivially modified for extensions such as the introduction of explanatory variables as well as the multivariate approach (Jacquier *et al.*, 1995). Furthermore,

increasing the number of parameters imposes a big computational cost to the MCMC procedure. For example, Jacquier *et al.* (1994) always remove the AR(1) component and the monthly effects in their empirical applications in order to decrease the number of parameters.

Alternatively, Danielsson and Richard (1993) proposes the simulated maximum likelihood (SML) estimator as a general method for estimation of dynamic latent models, and Danielsson (1994) applied it to the ARSV model. The likelihood in (14) is evaluated by simulation specifying a density $\rho(\theta)$ whose expectation equals $f(Y_T|\theta)$. By taking N draws from $\rho(\theta)$, a natural estimator of $f(Y_T|\theta)$ is given by

$$\hat{f}(Y_T|\theta) = \frac{1}{N} \sum_{i=1}^{N} \rho^{(i)}(\theta) \tag{15}$$

Once the likelihood is evaluated by simulation, estimates are obtained using a derivative-free optimizer. The SML estimator has the same asymptotic distribution as the ML estimator. Danielsson (1994) shows that SML and JPR have similar root mean square error (RMSE). However, his conclusions are based on just one experiment with $\phi = 0.9$ and $\sigma_\eta = 0.363$ and $T = 2000$.

With respect to the solution to the smoothing problem, Liesenfeld and Richard (2003) propose a procedure to estimate the underlying volatilities.

The SML estimator has been implemented to the empirical analysis of financial returns by, for example, Liesenfeld (1998), Liesenfeld and Jung (2000) and Liesenfeld (2001). Liesenfeld and Richard (2003) extends the SML estimator to SV models with non-Normal latent variables and errors and to multivariate models [see also Danielsson (1994, 1998) for leverage effect and multivariate extensions].

However, the SML procedure suffers from several drawbacks. First, the way in which the integral is evaluated is not direct, so that it is hard to measure the accuracy of the proposed approximation. In fact, the likelihood can only be exactly evaluated for $\phi = 0$ and for the rest of parameter values is not available. To solve this problem, Jacquier *et al.* (1994) propose to nest the SML procedure in a wider Bayesian framework using Monte Carlo methods of numerical integration. However, this extension is not yet developed. On top of that, Shephard (2000) shows that typical importance samplers for the ARSV model may not posses a variance and, consequently, not obey a standard central limit theorem.

Alternatively, Fridman and Harris (1998) propose a direct ML estimation method that calculates the likelihood function directly by means of the recursive numerical integration procedure suggested by Kitagawa (1987) for non-Gaussian-filtering problems. This method can be considered as an extended Kalman filter. Through a small simulation experiment, they conclude that the direct ML procedure performs better than QML and GMM, and similar to SML, the simulated expectation maximization (SEM) algorithm by Kim and Shephard (1994) and the JPR procedure. The smoothed sequence of volatilities can be obtained with an

additional forward-recursion step in the extended filter. They apply their technique to the S&P 500 index considering both Gaussian and Student-T errors.

Finally, Watanabe (1999) proposes a non-linear-filtering natural extension of QML. He obtains the exact form of the likelihood by means of a non-linear filter that makes use of the conditional probability density functions of the log-volatility and the observed series.[2] By means of a small Monte Carlo comparative study with the same design proposed by JPR, he shows that the efficiency of the non-linear-filtering maximum likelihood (NFML) procedure is close to the SML and JPR procedures. He also proposes a smoothing algorithm to estimate the volatilities that is shown to be superior to the standard smoothing solution used in QML. The NFML procedure can be used in models with a linear structure in the mean where the errors are assumed to be normal.

However, one of the main drawbacks of methods based on an extended Kalman filter [as Fridman and Harris (1998) or Watanabe (1999)] is their slow computational convergence. These procedures also involve choosing a priori a fixed grid over which the process will be integrated, and the optimal grid may not exist (Sandmann and Koopman, 1998).

3.1.3 *Estimation Procedures by Means of an Auxiliary Model*

SV models are easy to simulate, although they are difficult to estimate. The methods described in this subsection choose an auxiliary model that is easy to estimate. There are two main methods proposed within this group: Indirect Inference and Efficient Method of Moments (EMMs).

The indirect inference estimator proposed by Gourieroux *et al.* (1993) is given by

$$\tilde{\theta}_T^H = \underset{\theta \,\in\, \Theta}{\text{argmin}} \left(\hat{\beta}_T - \tilde{\beta}_{TH}(\theta) \right)' \hat{\Omega}_T \left(\hat{\beta}_T - \tilde{\beta}_{TH}(\theta) \right) \qquad (16)$$

where $\hat{\beta}_T$ is obtained by maximizing an auxiliary criterion from the auxiliary model $Q_T(\beta, \mathbf{y}_t)$, $\tilde{\beta}_{TH}(\theta)$ is an estimate of the binding function obtained by maximizing $Q_{TH}(\beta, \mathbf{y}_{TH}(\theta))$ and $\mathbf{y}_{TH}(\theta) = (y_1, \ldots, y_{tH})$ is a vector of simulated observations from the SV model. Gourieroux *et al.* (1993) propose the quasi-likelihood function of Harvey *et al.* (1994) as an auxiliary criterion to estimate a continuous time SV process (Pastorello *et al.*, 1994). Several alternative models have been proposed as auxiliary models. For example, Engle and Lee (1996) and Calzolari *et al.* (2001) use GARCH models, Monfardini (1998) uses AR(m) and ARMA(1,1) models, van der Sluis (1997) uses an EGARCH model and Fiorentini *et al.* (2002) use a non-linear asymmetric GARCH (NAGARCH) model.

The EMM approach, introduced by Bansal *et al.* (1993, 1995) and Gallant and Tauchen (1996), is similar but is based on score calibrating the criterion function. EMM improves the Indirect Inference approach in computational terms, as there is no need to refit the score to each simulated realization, while in the Indirect Inference procedure the binding function should be computed at each step.

The EMM procedure has been implemented by Engle (1994), Ghysels and Jasiak (1996), Andersen and Lund (1997), Chernov and Ghysels (2000), Dai and Singleton (2001) and Chernov *et al.* (2003), among others, to estimate SV models in continuous time. Gallant *et al.* (1997) and Jiang and van der Sluis (1998, 2000) use EMM to estimate discrete time univariate and multivariate SV models.

Gourieroux *et al.* (1993) show that the Indirect Inference and EMM estimators are asymptotically equivalent. Both estimators are consistent and asymptotically normal. Furthermore, Tauchen (1997) and Gallant and Long (1997) show that the estimated covariance matrix of the EMM estimator approaches that of ML, as the score generator approaches the true conditional density.

With respect to the finite sample properties of Indirect Inference and EMM estimators, Monfardini (1998) shows that the former estimator performs well in finite samples, but JPR and the SEM of Kim and Shephard (1994) procedures are more efficient. Andersen *et al.* (1999) observe that the EMM estimator is more efficient than GMM, although it is not as efficient as JPR. They also state that the efficiency of the method for finite samples strongly depends on the choice of the auxiliary model [see also Calzolari *et al.* (2000) for similar Monte Carlo results on a related estimator].

Gallant and Tauchen (1998) show how to obtain estimates of the underlying volatilities after the parameters have been estimated by EMM.

One of the main drawbacks of the Indirect Inference and EMM estimators is that they are very expensive in computational terms.

Table 1 summarizes the properties of the estimators described in this subsection.

3.2 *Methods Based on* $\log(y_t^2)$

3.2.1 *Quasi-Maximum Likelihood*

The QML estimator, proposed independently by Nelson (1988) and Harvey *et al.* (1994), is based on linearizing the SV model by taking logarithms of squares as in model (7) and (8). Treating ξ_t as if it were Gaussian, the Kalman filter can be applied in order to obtain the quasi-likelihood function of $\log(y_t^2)$ which, ignoring constants, is given by

$$\log L\big(\log(y^2)|\theta\big) = -\frac{1}{2}\sum_{t=1}^{T}\log F_t - \frac{1}{2}\sum_{t=1}^{T}\frac{\nu_t^2}{F_t} \qquad (17)$$

where ν_t is the one-step-ahead prediction error of $\log(y_t^2)$ and F_t is the corresponding mean-squared error. Ruiz (1994) shows that the QML estimator is consistent and asymptotically normal. However, the QML procedure is inefficient, as the method does not rely on the exact likelihood of $\log(y_t^2)$. Note that approximating the density of ξ_t by a normal density instead of using the true $\log(\chi_1^2)$ density could be rather inappropriate (see Figure 1 for a comparison of

Table 1. Summary of Estimators of Parameters of ARSV(1) Model Based on Series y_t.

Method	Authors	Asymptotic Properties	Finite Sample Experiments	Advantages/Limitations
GMM	Taylor (1986) Melino and Turnbull (1990)	Consistent and asymptotically normal (Hansen, 1982)	Efficiency depends on moments Less efficient than MCMC	Not computationally intensive h_t stationary Bad results for ϕ close to 1
SMM	Duffie and Singleton (DS) 1993	Consistent and asymptotically normal (DS 1993)	Not available	Analytical moments not used
Importance sampling	Geweke (1994)	Simulation based	Not available	Bad results if T is big or dimension of h_t bigger than 1 (a); (e)
MCMC	Jacquier et al. (1994)	Simulation based	More efficient than GMM or QML	Computationally intensive non-trivial extensions (b); (c); (d); (e)
SML	Danielsson and Richard (1993) Danielsson (1994)	As ML estimators	Not available	(b); (c); (e)
ML	Fridman and Harris (1998)	As ML estimators	Outperforms: GMM, QML Similar: SML, MCMC, SEM	Slow convergence (b); (e)
NFML	Watanabe (1999)	As ML estimators	Outperforms GMM, QML Similar to MCMC and SML	Slow convergence (a); (e)
Indirect inference	Gourieroux et al. (1993) (GMR)	Consistent and asymptotically normal (GMR)	Outperformed by MCMC and SEM	(b)
EMM	Gallant and Tauchen (1996)	Consistent and asymptotically normal (GMR)	Outperforms GMM Outperformed by MCMC	

Notes Last column codes are the following: (a), dynamics on the conditional mean; (b), fat-tailed distributed errors; (c), multivariate extension; (d), leverage effect; (e), estimation of the latent volatilities.
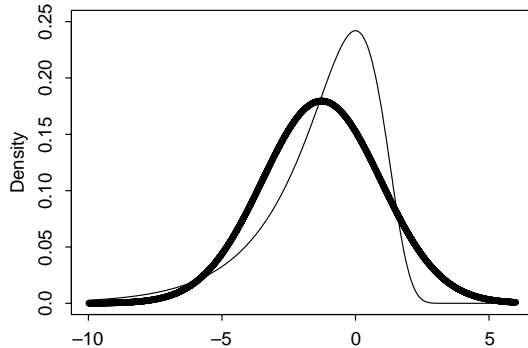
Figure 1. The $\log(\chi_1^2)$ Density (Solid Line) and its Corresponding Normal Approximation (Thick Solid Line).

both densities). The effects of this approximation depend on the true parameter values and are worse, as the variance of the volatility equation σ_η^2 decreases.

On the other hand, an inlier problem often arises when dealing with the log-squared transformation. When returns, y_t, are very close to zero, the log-squared transformation yields large negative numbers. In the extreme case, if the return is equal to 0, the log-squared transformation is not defined. In Figure 1, large negative values in the distribution of $\log(\chi_1^2)$ reflect the presence of those inliers. To solve this problem, Fuller (1996) proposes the following modification of the log-squared transformation

$$y_t^* = \log(y_t^2 + \tau s^2) - \frac{\tau s^2}{y_t^2 + \tau s^2} \qquad (18)$$

where s^2 is the sample variance of y_t and τ is a small constant. In several studies, this constant has been set equal to 0.02 (Fuller, 1996; Breidt and Carriquiry, 1996; Bollerslev and Wright, 2001).

The finite sample properties of the QML estimator have been analysed by Ruiz (1994) who shows that the bias of the estimator of σ_η^2 increases when σ_η^2 decreases. Jacquier *et al.* (1994) also find that QML has inadequate finite sample properties when the persistence is high and σ_η^2 is small. Furthermore, they also remark that, in this case, there is a huge degradation in the filtering performance in terms of the RMSE. However, Sandmann and Koopman (1998) and Breidt and Carriquiry (1996) state that although QML is inefficient, it is not as bad as Jacquier *et al.* (1994) show. These authors suggest that the bad results of QML in Jacquier *et al.* (1994) may be due to an inefficient implementation of the procedure (poor starting values, different convergence criteria, etc.). On the other hand, Andersen and Sørensen (1996) show that the QML estimator dominates the GMM

estimator for models with a high degree of persistence. Deo (2002) provides theoretical intuition for this finding.

Once the parameters have been estimated, Harvey *et al.* (1994) propose to implement a smoothing algorithm based on the Kalman filter to obtain estimates of the underlying log-volatilities. However, the Kalman filter yields minimum mean square linear estimators (MMSLEs) of h_t rather than minimum mean square estimators (MMSEs).

Despite the limitations of the method, QML procedure is very flexible and has been extensively implemented for the empirical analysis of financial returns[3] [see, for example, Hwang and Satchell (2000) and Yu (2002a)]. Furthermore, several generalizations of the method have been proposed. For instance, the QML estimator can be directly implemented to estimate models with heavy-tailed errors and can also be easily extended to models with explanatory variables or other ARMA models for the log-volatility. Furthermore, the multivariate generalization is straightforward in this context (Harvey *et al.*, 1994; Lien and Wilson, 2001; McMillan, 2001). Missing or irregularly spaced observations can also be easily handled. Finally, Harvey and Shephard (1996) show that, after imposing certain distributional assumptions about ε_t and η_t, the QML estimator of the SV model with leverage effect is consistent and asymptotically normal.

Finally, Alizadeh *et al.* (2002) propose a QML estimation procedure based on using the range as a proxy of the volatility. The range is defined as the difference between the highest and lowest log security prices over a fixed sampling interval. The new procedure is applied to five exchange rates.

3.2.2 *Other Methods Based on Linearization*

In this subsection, we describe methods that try to approximate the likelihood of $\log(y_t^2)$. Originally, Kim and Shephard (1994) proposed a simulated expectation maximization (SEM) algorithm using a mixture of seven normal distributions to match the first four moments of $\log(\varepsilon_t^2)$. Later, Mahieu and Schotman (1998) propose a more flexible mixture in order to accommodate a wider range of shapes of $\log(\varepsilon_t^2)$. The main advantage of approximating the distribution of $\log(\varepsilon_t^2)$ by mixtures of normals is that, conditional on the mixture component, the state space model in (7) and (8) is Gaussian. In addition, the use of mixtures makes the procedure more robust than QML to the inlier problem. A multimove Gibbs sampling technique that extends the usual Gaussian Kalman filter can then be applied (Carter and Kohn, 1994; Shephard, 1994).

The performance of the SEM procedure for finite samples has been studied by Fridman and Harris (1998), who show that this procedure is similar to MCMC and SML in terms of efficiency.

Later, Kim *et al.* (1998) (KSC) propose a procedure that nests and improves several aspects of the SEM estimator. Specifically, they propose a MCMC algorithm that samples all the unobserved volatilities simultaneously by means of an approximating offset mixture of normal model, together with an importance

reweightening procedure to correct the linearization error. The KSC procedure provides efficient inferences, likelihood evaluation, filtered volatility estimates, diagnostics for model failure and computation of statistics for comparing non-nested volatility models. However, its finite sample properties have not been compared with other estimators [see Steel (1998) for an alternative MCMC algorithm for $\log(y_t^2)$].

The generalization of the KSC method to Student-t errors was proposed by Chib *et al.* (2002a) who also include a generalization of the method for the SV model with jumps.

Alternatively, Sandmann and Koopman (1998) propose the Monte Carlo Likelihood (MCL) procedure that approximates the likelihood function of $\log(y_t^2)$ by a Gaussian part constructed via the Kalman filter plus a correction for departures from the Gaussian assumption relative to the true unknown model as follows

$$\log L\big(Y_T^*|\theta\big) = \log L_G(Y_T^*|\theta) + \log E_G\left(\frac{f_{\log \chi_1^2}(\xi|\theta)}{f_G(\xi|\theta)}\right) \tag{19}$$

where $Y_T^* = (\log(y_1^2), \ldots, \log(y_T^2))$, $L_G(Y_T^*|\theta)$ is the Gaussian likelihood function, $f_{\log \chi_1^2}(\xi|\theta)$ is the true $\log(\chi_1^2)$ density, $f_G(\xi|\theta)$ is the importance density corresponding to the approximating Gaussian model and E_G refers to the expectation with this density. The MCL is based on estimating model (7) and (8) assuming conditional Gaussianity and computing $\log L_G(Y_T^*|\theta)$ and the corresponding estimates of $\hat{\xi}_t$. Then, using Shephard and Pitt (1997) and Durbin and Koopman (1997), N importance samples from the importance density are generated, $\tilde{\xi}_t^{(i)}$, $i=1,\ldots,N$ together with the quantities $\omega(\tilde{\xi}_t^{(i)}) = f_{\log \chi_1^2}(\tilde{\xi}_t^{(i)}|\theta)/f_G(\tilde{\xi}_t^{(i)}|\theta)$. Finally, the sample mean and variance through all importance samples of $\omega(\tilde{\xi}_t^{(i)})$, denoted by $\overline{\omega}$ and s_w^2, respectively, are computed. The unbiased estimate of $\log L(Y_T^*|\theta)$ in (19) is given by

$$\log \widehat{L\big(Y_T^*|\theta\big)} = \log L_G\big(Y_T^*|\theta\big) + \log \overline{\omega} + \frac{s_w^2}{2N\overline{\omega}^2} \tag{20}$$

Then, MCL estimates of the parameters are obtained by numerical optimization of (20). Sandmann and Koopman (1998) recommend a small number of draws, $N=5$, to obtain similar efficiency as compared with MCMC methods. The MCL procedure also generates simultaneously estimates of the latent volatilities. The Kalman filter smoother applied to the approximating Gaussian model (7) and (8) effectively computes the posterior mode estimates of the volatility. If the posterior mean is required, it is possible to use the following computationally efficient algorithm of Durbin and Koopman (2000):

$$\log\left(\sigma_{t/T}^2\right) = \log(y_t^2) - \sum_{i=1}^{N} f_t \tilde{\xi}_t^{(i)}$$

where $f_t = \frac{\omega(\xi_t^{(i)})}{N\bar{\omega}}$. Notice that the smoothness of MCL estimates of the underlying log-volatilities can be controlled by the number of replications in the importance sampler.

The MCL estimation procedure can be easily generalized to heavy-tailed errors, explanatory variables or to the inlier problem. The multivariate case is not yet available. Although Sandmann and Koopman (1998) also mention that the MCL can be implemented in models with correlated ε_t and η_t and with stochastic seasonal components, these extensions have not yet been developed in the literature.

Brandt and Kang (2004) propose an interesting application of the MCL procedure for monthly returns on the CRSP index and use MCL to fit a VAR model where the first equation describes the dynamics of the conditional mean and the second equation defines the variance as a SV model.

The MCL procedure has also been generalized to the SVM model by Koopman and Uspensky (2002) who estimate the underlying volatilities by means of the particle-filtering technique of Pitt and Shephard (1999a).

Recently, Singleton (2001) and Knight *et al.* (2002) have proposed a new estimation procedure based on the empirical characteristic function (ECF). As $\log(y_t^2)$ is the convolution of an AR(1) and an independent $\log(\chi_1^2)$ process, there is a closed form expression for the characteristic function, so that the model is fully and uniquely parameterized by it. Knight and Yu (2002) establishes the strong consistency and asymptotic normality for the ECF estimators with a general weight function.

Table 2 summarizes the main estimators of the parameters of the ARSV(1) model described in this subsection, together with their asymptotic properties and their main advantages.

3.3 *Illustration with Simulated Data*

The objective of this section is to show how the main estimators described in this article can be implemented to real data and to illustrate whether the estimates of the parameters and volatilities obtained are different. Notice that a proper comparison of the properties of the alternative estimators requires a Monte Carlo study that is beyond the objectives of this survey. We simulate one series of size $T = 5000$ by each of the two following ARSV models: $\phi = 0.95$, $\sigma_\eta^2 = 0.05$ and $\sigma_* = 1.0$ (M1) and $\phi = 0.98$, $\sigma_\eta^2 = 0.02$ and $\sigma_* = 1.0$ (M2) The parameters have been selected to represent values often found in empirical applications of daily returns. The estimation is carried out using the whole series and using also the first 500 and 1500 observations by GMM, JPR, QML, MCL and KSC. With respect to GMM, estimates of the covariance of the differences between the sample and populational moments have been obtained as in Melino and Turnbull (1990).[4] We have not obtained the original code of the JPR estimator and, consequently, we have implemented our own code following Wong (2002) with the number of iterations suggested by JPR, i.e. 1500 burn-in iterations and 2500 iterations.[5] The convergence of the resulting algorithm is extremely slow.[6] The

Table 2. Summary of Estimators of Parameters of ARSV(1) Model Based on Series $\log(y_t^2)$.

Method	Authors	Asymptotic Properties	Finite Sample Experiments	Advantages and Limitations
QML	Harvey et al. (1994)	Consistent and asymptotically normal (Ruiz, 1994)	Inefficient	Not computationally intensive (b); (c); (d); (e)
SEM	Kim and Shephard (1994)	Simulation based	Similar to ML methods	
KSC	Kim et al. (1998)	Simulation based	Not available	(b); (e)
MCL	Sandmann and Koopman (1998)	As ML estimators	Similar to MCMC	
ECF	Singleton (2001) Knight et al. (2002)	Strong consistency and asymptotically normal (Knight and Yu, 2002)	Not available	(a); (b); (e) (d)

Notes Last column codes are the following (a), dynamics on the conditional mean; (b), fat-tailed distributed errors; (c), multivariate extension; (d), leverage effect; (e), estimation of the latent volatilities.

MCL procedure is implemented using the library SsfPack 2.3 of Koopman, Shephard and Doornik (1999).[7] The program employed for estimation is named sv_mcl_est.ox and can be downloaded from the website of Siem Koopman.[8] The KSC procedure to estimate the parameters is implemented in BUGS following the approach by Meyer and Yu (2000)[9] with 10,000 iterations and 1000 burn-in iterations. The underlying volatilities have been estimated using the software package SVPack 2.1 by Neil Shephard.[10]

Estimates of the scale parameter, σ_*, are obtained in different ways depending on the estimation procedure. In the QML case, $\hat{\sigma}_*$ is the sample SD of the observations standardized by the smoothed estimates of the volatility [see Harvey and Shephard (1993) for details on this standardization]. The approximation of the asymptotic variance of $\hat{\sigma}_*$ is given by $Var(\hat{\sigma}_*^2) = (\pi^2/2)(\hat{\sigma}_*^2)^2/T$. In MCL and JPR, the corresponding algorithms obtain estimates of $\gamma = (1 - \phi)\log(\sigma_*^2)$, hence estimates of σ_* and its corresponding SD are obtained after transformation. Finally, KSC estimates directly σ_*.

Table 3, that reports the estimation results for model M1, shows that when $T = 500$, the GMM estimates of ϕ and σ_η^2 underestimate the true values and suggests that the volatility is nearly constant over time. Furthermore, the SDs of these estimates are clearly greater than those for any of the other methods considered. Finally, we would like to point out that, for this moderate sample size and the particular series generated, we encounter important difficulties for the convergence of the GMM estimation algorithm. The estimates of ϕ and σ_η^2 obtained by the other methods are rather similar. With respect to their SEs, the biggest corresponds to the QML estimator and the smallest to the KSC. Looking at the estimates of the scale parameter, it can be observed that all of them are very similar. However, in this case, the smallest SE corresponds to the GMM estimator and the biggest to the KSC estimator. Finally, for the larger sample sizes, $T = 1500$ and 5000, the estimates of all the parameters are similar independently of the estimator used. The main differences between the estimators arise in the SEs, with the GMM SEs of $\widehat{\phi}$ and $\widehat{\sigma}_\eta^2$ being rather big compared with the other estimators. Notice that the results for QML and KSC are remarkably similar. Once more, the smallest SEs of the estimates of σ_* are obtained using the GMM and QML estimators.

The results for the M2 model are reported in Table 4 and are very similar to the ones reported in Table 3. Although these results have just scratched the surface of the problem and, as we mentioned before, they have just an illustrative purpose, it seems that the estimator of the scale parameter based on the corresponding sample moment has good properties. Therefore, as suggested by Harvey and Shephard (1993), it seems that to standardize the original observations using the estimated marginal SD before the ARSV model is estimated and then to estimate the scale parameter by the sample SD of the observations standardized by the estimated volatilities could be a reasonable strategy. Furthermore, for the sample sizes usually encountered in empirical applications, it seems that, with the exception of the GMM estimator, the estimates of the parameters obtained by the alternative estimators considered in this example are rather similar.

Table 3. Estimation Results for One Series Generated by ARSV(1).

	JPR	GMM	QML	MCL	KSC
$T = 500$					
$\phi = 0.95$	0.9510 (0.0357)	0.7376 (1.3524)	0.9267 (0.0557)	0.9429 (0.0300)	0.9578 (0.0216)
$\sigma_\eta^2 = 0.05$	0.0461 (0.0465)	0.0168 (0.3967)	0.0656 (0.0664)	0.0571 (0.0315)	0.0398 (0.0188)
$\sigma_* = 1.0$	0.8480 (0.0855)	0.9837 (0.0304)	0.9275 (0.0461)	0.9318 (0.0920)	0.9865 (0.1301)
$T = 1500$					
$\phi = 0.95$	0.9652 (0.0310)	0.9622 (0.1491)	0.9511 (0.0254)	0.9496 (0.0174)	0.9576 (0.0134)
$\sigma_\eta^2 = 0.05$	0.0456 (0.0296)	0.0501 (0.0749)	0.0496 (0.0204)	0.0570 (0.0194)	0.0459 (0.0141)
$\sigma_* = 1.0$	0.9754 (0.0770)	1.0013 (0.0035)	1.0973 (0.0314)	1.0694 (0.0784)	1.0830 (0.0770)
$T = 5000$					
$\phi = 0.95$	0.9612 (0.0140)	0.9510 (0.0276)	0.9527 (0.0112)	0.9622 (0.0073)	0.9632 (0.0069)
$\sigma_\eta^2 = 0.05$	0.0406 (0.0121)	0.0541 (0.0163)	0.0443 (0.0130)	0.0371 (0.0072)	0.0363 (0.0072)
$\sigma_* = 1.0$	0.9887 (0.0372)	0.9993 (0.0074)	1.0136 (0.0159)	1.0037 (0.0382)	1.0080 (0.0394)

Note Standard errors in parenthesis.

Table 4. Estimation Results for One Series Generated by ARSV(1).

	JPR	GMM	QML	MCL	KSC
$T = 500$					
$\phi = 0.98$	0.9700 (0.0318)	0.9836 (2.5064)	0.9535 (0.0474)	0.9718 (0.0171)	0.9742 (0.0156)
$\sigma_\eta^2 = 0.02$	0.0111 (0.0265)	0.0084 (1.3450)	0.0235 (0.0322)	0.0207 (0.0121)	0.0186 (0.0097)
$\sigma_* = 1.0$	0.9068 (0.0759)	0.9979 (0.3169)	0.8744 (0.0434)	0.8784 (0.1011)	0.9348 (0.1307)
$T = 1500$					
$\phi = 0.98$	0.9871 (0.0197)	0.9909 (0.1586)	0.9890 (0.0056)	0.9864 (0.0058)	0.9861 (0.0056)
$\sigma_\eta^2 = 0.02$	0.0186 (0.0162)	0.0066 (0.0301)	0.0112 (0.0060)	0.0158 (0.0055)	0.0173 (0.0047)
$\sigma_* = 1.0$	0.9596 (0.0444)	0.9997 (0.0118)	1.1276 (0.0323)	1.0908 (0.1511)	1.0820 (0.1253)
$T = 5000$					
$\phi = 0.98$	0.9891 (0.0085)	0.9759 (0.1284)	0.9905 (0.0027)	0.9895 (0.0026)	0.9895 (0.0025)
$\sigma_\eta^2 = 0.02$	0.0184 (0.0074)	0.0182 (0.0227)	0.0102 (0.0029)	0.0121 (0.0022)	0.0124 (0.0019)
$\sigma_* = 1.0$	0.9661 (0.0217)	0.9978 (0.0035)	1.0089 (0.0158)	1.0000 (0.0369)	1.0070 (0.0739)

Note Standard errors in parenthesis.

The final objective when fitting a SV model to a series of financial returns is usually to obtain estimates of the underlying volatilities. Although the parameters estimated by the different methods are not very different, we will see whether the estimated volatilities are different. Figure 2 plots the corresponding volatilities estimated by the QML, JPR, KSC and MCL filters, together with the underlying simulated volatilities for the M1 model and $T = 1500$. Notice that the KSC estimates are too volatile while the QML estimates are too smooth when compared with the true underlying volatilities. The corresponding RMSE are 0.32 and 0.29, respectively. On the other hand, the JPR and MCL estimates are rather similar and close to the true volatilities. In both cases, the RMSE is 0.27. Finally, Figure 3 plots the estimated volatilities obtained using the QML, JPR and MCL filters with the parameters fixed at the GMM estimates together with the simulated volatilities. Once more, the shapes of the JPR and MCL estimates are quite similar with RMSE of 0.27. As before, the QML estimates have a smoother shape, and the RMSE is slightly bigger with a value of 0.29. Comparing Figs 2 and 3, it is possible to observe that, for the particular series generated in this example, the differences between the estimated sequences of volatilities are not so different when the alternative filters are run using the same GMM estimates of the parameters as when they are run with different estimates. Taking into account that, the parameters estimated by the alternative methods are very similar, our results seem to suggest that small differences in the parameters can generate relatively large differences between the sequences of volatilities estimated by the alternative filters.

4. Estimation Methods for LMSV Models

In this section, we describe estimators proposed for the parameter of LMSV. Most of these methods are based on the spectral representation of the model.

Harvey (1998) and Breidt *et al.* (1998) propose to estimate the parameters of the LMSV model in (11) by QML, maximizing the Whittle discrete approximation to the Gaussian likelihood function of $\log(y_t^2)$ in the frequency domain, given by

$$\tilde{L}(\theta|y) = -\frac{1}{2T} \sum_{j=1}^{T-1} \left[\log\ f(\lambda_j, \theta) + \frac{I(\lambda_j)}{f(\lambda_j, \theta)} \right] \tag{21}$$

where $f(\lambda_j, \theta)$ is the spectral density, $\lambda_j = \frac{2\pi j}{T}$ are the corresponding frequencies and $I(\lambda_j)$ is the sample spectrum. Breidt *et al.* (1998) demonstrate that the QML estimator is strongly consistent. Furthermore, Deo (2002) proves the asymptotic normality of the QML estimator obtained maximizing the time domain likelihood and conjectures that this result may also hold for the Whittle estimator. With respect to its finite sample properties, Breidt *et al.* (1998) conclude that they are satisfactory. However, their parameter design is not very realistic and Pérez and Ruiz (2001) extend their analysis and conclude that, when parameters are close to non-stationarity ($d \approx 0.5$) and/or to homoscedasticity ($\sigma_\eta^2 \approx 0$), the properties of
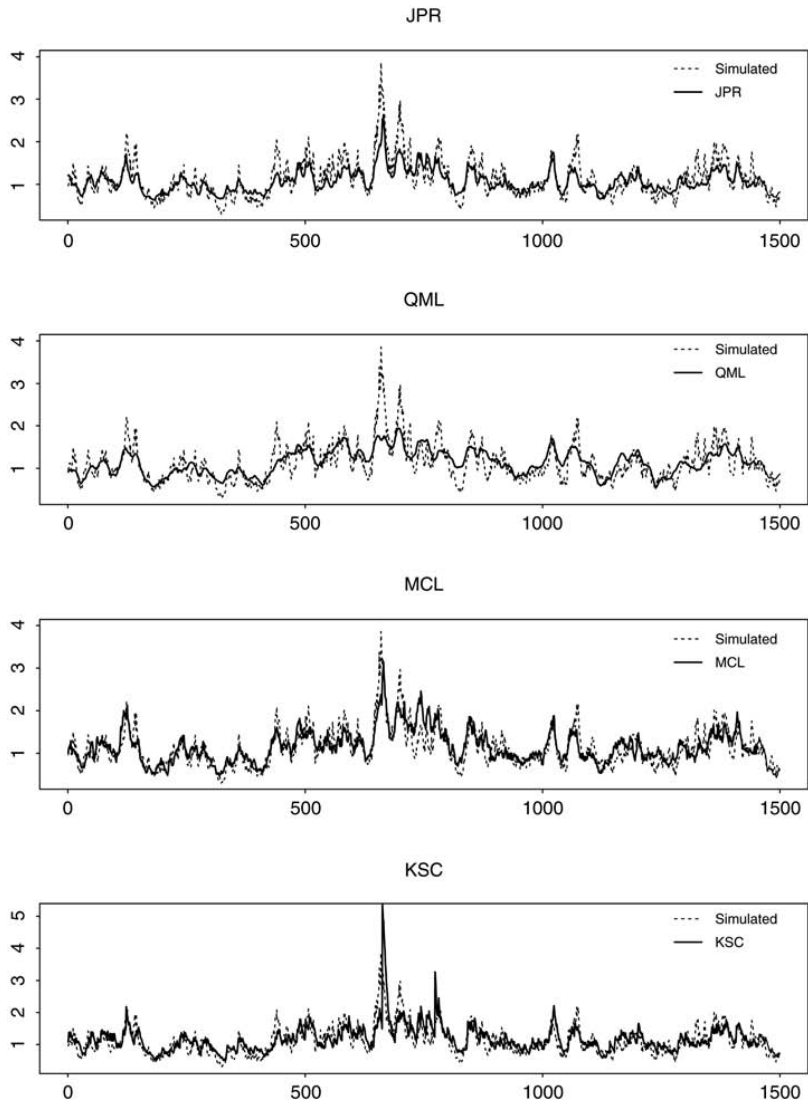
Figure 2. Estimated Volatilities of a Simulated Series with $\phi = 0.95$ and $\sigma_\eta^2 = 0.05$ and $T = 1500$.

Volatilities - simulated series
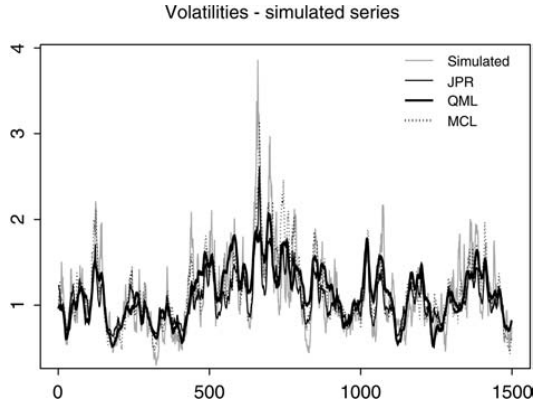
Figure 3. Estimated Volatilities of a Simulated Series with $\phi = 0.95$ and $\sigma_\eta^2 = 0.05$ Estimated by JPR, QML and MCL Using Estimates Obtained by GMM. $T = 1500$.

the QML estimator are very poor, so that huge sample sizes are needed to obtain reliable inferences.[11] The Whittle estimator has been implemented by Islas-Camargo and Venegas-Martínez (2003) to estimate the LMSV model fitted to several indexes of Latin American Stock Exchanges.

Harvey (1998) suggests an algorithm to obtain the smoothed estimates of the volatility. The empirical implementation of the algorithm is developed in Pérez (2000) and applied to a real series in Pérez and Ruiz (2001).

Estimation of LMSV models has also been carried out by means of Bayesian procedures. Hsu and Breidt (1997) obtain the posterior distribution of the parameters and the smoothed estimates of the volatilities of LMSV processes by means of a Gibbs sampling algorithm. So (1999) develops a new algorithm based on the state space formulation of Gaussian time series models with additive noise where full Bayesian inference is implemented through MCMC techniques. This algorithm can be applied to model outliers in long-memory time series and LMSV models [see So (2002) for an empirical application to S&P500 returns]. Finally, Chan and Petris (2000) propose a Bayesian approach to perform inference in the time domain based on the truncated likelihood method where the LMSV model is expressed as a linear state space model. Nevertheless, there is no finite sample analysis, so that the behaviour of this procedure compared with other methods is unknown. Finally, Jensen (2003) extends the KSC procedure to estimate the fractional parameter using a wavelet representation of the log-squared returns.

All the three Bayesian-based approaches are computationally intensive and not easy to generalize to more complicated models, for example, when the log-volatility equation has an ARMA component.

Deo and Hurvich (2001) suggest a semiparametric estimator of the parameter d based on the GPH estimator of Geweke and Porter-Hudak (1983) for ARFIMA models and derive its asymptotic bias and variance. They assume that $\phi = 0$ in model (11) and show that, under certain conditions, the corresponding asymptotic distribution is normal, although the convergence rate is $\sqrt{m}$, where m is the number of Fourier frequencies used in the corresponding log-periodogram regression. A small Monte Carlo experiment is carried out for a unique sample size of $T = 6144$. The asymptotic normality of the GPH estimator relies on the Gaussianity of the log-volatility process. Bollerslev and Wright (2000) show that the properties of the GPH estimator of d are closely related to the sampling frequency of the data and propose new estimators that aggregate the information in very high frequency returns. Alternatively, Arteche (2004) proposes a local Whittle estimator of the long-memory parameter that relaxes this assumption. This estimator is consistent and asymptotically normal and more efficient than the estimator based on the log-periodogram regression. However, the local Whittle estimator is also affected by the large and negative biases often found in this context (Crato & Ray, 2002). Finally, it is interesting to mention that the SV model considered by Arteche (2004) incorporates persistent stochastic seasonality in the volatility and therefore could be implemented to intradaily data.

Finally, Wright (1999) proposes a GMM estimator of LMSV models and demonstrates its consistency and asymptotic normality, provided that $1/2 < d < 1/4$. However, numerous studies have found that usually the estimates of d for high-frequency returns are between 0.3 and 0.47 (Andersen *et al.*, 2001). Therefore, the case $d \geq 1/4$ is most interesting from an empirical point of view. Monte Carlo experiments comparing the GMM and the frequency domain estimators show that the asymptotic distribution of the GMM estimator is not an adequate approximation for finite samples, even for the biggest sample size considered, $T = 4000$. Furthermore, although both estimators have similar SEs, the biases are bigger for GMM. Deo (2002) shows that, with the moment conditions that have been commonly used, the rate of convergence of the GMM estimator is $T^{1/2 \ d}$ Alternatively, he proposes a new set of moment conditions based on the linear transformation $\log(y_t^2)$ and shows that, in this case, the GMM estimator is $\sqrt{T}$ consistent and asymptotically normal.

Table 5 provides a summary of the main characteristics of the estimators proposed for LMSV models.

5. Empirical Application

In this section, the ARSV(1) model is fitted to daily observations of the S&P 500 stock price index to illustrate with real data which are the differences between the estimated parameters and volatilities depending on the estimator used. The series is observed daily from 19 February 1997 to 15 February 2002 and $T = 1303$.[12] The prices, p_t, are transformed into returns, r_t, in the usual way, i.e. $r_t = 100 \cdot \log(p_t/p_{t-1})$ and centred around the sample mean. Figure 4 plots the series of returns and the corresponding autocorrelations of the squared returns.

Table 5. Summary of Estimators of Parameters of LMSV Model.

Method	Authors	Asymptotic Properties	Finite Sample Experiments	Advantages and Limitations
QML	Harvey (1998) Breidt *et al.* (1998)	Consistent	Poor properties when $d \sim 0.5$ and/or $\sigma_\eta^2 \approx 0$	Estimation of volatility
MCMC	Hsu and Breidt (1997) So (1999) Chan and Petris (2000)	Simulation based	Not available	Computationally intensive Not easy to generalize Estimation of volatility
Semi-parametric	Deo and Hurvich (2001)	Results only for $\hat d$ Asymptotically normal under certain conditions	Not exhaustive Monte Carlo design	
GMM	Wright (1999)	Consistent and asymptotically normal if $d < 1/4$	Similar to QML	Unknown asymptotics for $d > 1/4$
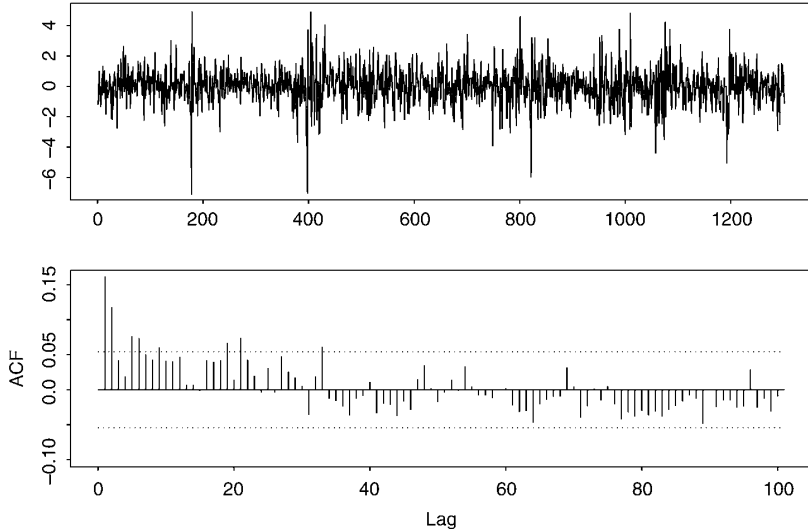
Figure 4. S&P 500 Stock Index Returns Observed daily from 19 February 1997 to 15 February 2002 and Correlogram of the Squared Returns.

Table 6 reports several sample moments of the daily and squared returns. As usual, returns are leptokurtic and uncorrelated, although not independent, given that their squares have significant autocorrelations. However, there is not evidence of long memory in squared returns.[13] Therefore, we fit the short memory ARSV model to the series r_t.

Table 7 shows the parameter estimates of the ARSV model obtained by the GMM, JPR, QML, MCL and KSC estimators. All the estimates of the persistence parameter, ϕ, are rather similar being the smallest, 0.93, obtained when the MCL estimator is used and the biggest, 0.96, when the parameters are estimated

Table 6. Summary Statistics of r_t.

	Mean	SD	SK	Kurtosis	Maximum	Minimum	Q (10)
r_t	0.000	1.249	0.232	0.813*	4.965	7.136	13.777
r_t^2	1.560	3.425	7.270	82.469	50.926	0.000	82.799*

	$\rho(1)$	$\rho(2)$	$\rho(3)$	$\rho(4)$	$\rho(5)$	$\rho(6)$	$\rho(7)$	$\rho(8)$	$\rho(9)$	$\rho(10)$
r_t	0.005	0.049	0.046	0.006	0.021	0.020	0.048	0.000	0.000	0.052
r_t^2	0.162*	0.118*	0.042	0.019	0.076*	0.074*	0.051	0.043	0.060*	0.041

Note: *Significant at the 5% level.

by GMM or JPR. However, the corresponding standard deviations are rather different. When the parameters are estimated by JPR, MCL or KSC, the standard deviations are approximately 0.02, but when they are estimated by GMM or QML, they are much bigger, 0.05 and 0.07, respectively. This result is in agreement with the finite sample results described in previous sections. With respect to the estimates of the variance of volatility, σ_η^2, there are important differences depending on the estimator used. For example, if the parameters are estimated by JPR or QML, the estimates are approximately, 0.01, and consequently, the evolution of the volatility is very smooth. However, when the parameters are estimated by GMM, MCL or KSC, the variance of the underlying volatilities is much bigger, 0.05. Notice that these differences on the estimates of the variance are also reflected on the corresponding estimates of the CV that go from 0.1164 to 1.0006 when the parameters are estimated by QML or GMM, respectively. In this example, the standard deviations of the estimates of σ_η^2 are rather similar. Finally, looking at the estimates of the scale parameter, σ_*, all of them are around 1. Once more, the main differences appear in the standard deviations although, in this case, the smallest corresponds to GMM and the biggest to JPR.

The estimates of the volatility obtained by the QML, JPR, MCL and KSC procedures are plotted in Figure 5. All four volatility estimates detect a similar dynamic evolution of the latent variable, although, as we have already observed with the simulated data, the QML procedure produces smoother estimates than MCL and KSC which are rather similar. Finally, we have obtained the estimates of the underlying volatilities using the QML, JPR and MCL filters with the parameters fixed at the GMM estimates. The results appear in Figure 6, where it can be observed that, once the parameters are fixed at the GMM estimates, all the filters give similar estimates of the volatility. This result, that has been also observed with the simulated data, is important because it shows that the differences between the estimated volatilities plotted in Figure 5 could be attributed mainly to differences among the estimated parameters and not to the filters used to estimate them.

6. Conclusions

In this article, the main estimation procedures of SV models have been revised. There are several methods that seem to match the benchmark efficiency established

Table 7. Empirical Estimates of ARSV Model for the S&P 500.

	ARSV				
	GMM	JPR	QML	MCL	KSC
$\hat{\phi}$	0.9602 (0.0479)	0.9596 (0.0203)	0.9401 (0.0699)	0.9288 (0.0249)	0.9392 (0.0237)
$\hat{\sigma}_\eta^2$	0.0541 (0.0219)	0.0172 (0.0196)	0.0128 (0.0222)	0.0499 (0.0190)	0.0405 (0.0168)
CV	1.0006	0.2427	0.1164	0.4381	0.4099
$\hat{\sigma}_*$	1.0005 (0.0157)	0.9673 (0.0639)	1.2051 (0.0371)	1.1248 (0.0542)	1.1260 (0.0619)
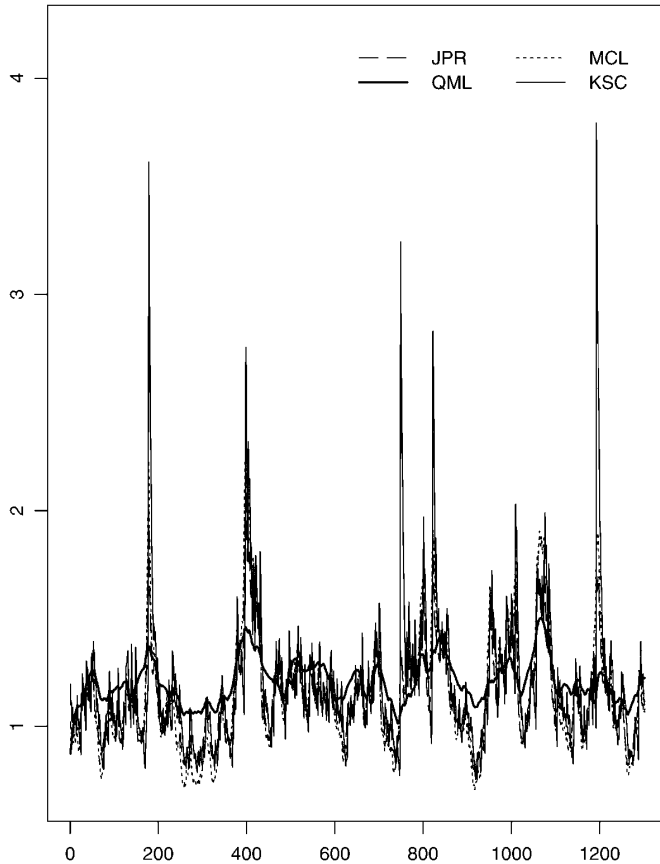
Figure 5. Estimated Volatilities of S&P 500 by JPR, QML, KSC and MCL.

by the MCMC procedure of Jacquier *et al.* (1994), like MCL, SML and ML of Fridman and Harris (1998) or the NFML approach of Watanabe (1999) being, at the same time, simpler from a computational point of view. On the other hand, there are very simple although not efficient methods as for example QML that can be easily implemented in real time series. Whether, for the sample sizes usually encountered in financial time series, the lost of efficiency compensates of using the more computationally demanding methods is still an open question. It is evident the need of a detailed comparative study of the finite sample properties of the
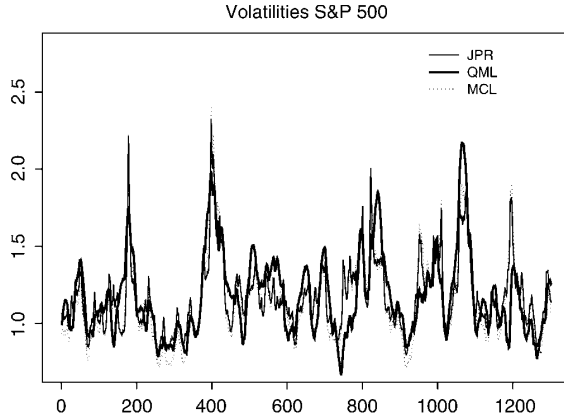
Figure 6. Estimated Volatilities of S&P 500 by JPR, QML and MCL Using Estimated Parameter Values Obtained by GMM.

main estimators of the parameters of SV models. In fact, some of the Monte Carlo experiments are only available for one set of parameter values.

With respect to results on the properties of the alternative procedures to estimate the latent volatilities, only Jacquier *et al.* (1994) compare volatility estimates calculated by MCMC with Kalman filter estimates obtained using as parameters for the filter the QML and GMM estimates. In this article, we carry out a very limited illustrative example comparing the estimates of volatility obtained by the alternative algorithms considered, using two simulated and one real time series. The estimates obtained by the simplest QML filter are the smoothest, followed by the JPR and MCL estimates that are similar. Finally, the estimates obtained by the KSC algorithm are the spikiest. Finally, we show that when the filters are run with the parameters fixed at the GMM estimates, all of them give similar results.

Acknowledgements

Notes

1. The QML estimator will be described in section 2.2.

2. This is not the first attempt to use non linear filters to estimate SV models; see, for example, Brigo and Hanzon (1998).
3. The QML procedure for SV models is already implemented in the software STAMP 6.0 by Koopman, Harvey, Doornik and Shephard, London: Timberlake consultants (2000).
4. Andersen and Sørensen (1996) suggest that better results could be obtained using alternative windows.
5. We are very grateful to Mike Wiper for his continuos help to develop this code.
6. Geweke (1994) proposes an alternative algorithm that uses an adaptive rejection sampling algorithm that could be faster.
7. More information is available at http://www.ssfpack.com.
8. We are very grateful to Siem Koopman for helpful suggestions using this code.
9. BUGS is available free of charge from http://www.mcr bsu.cam.ac.uk/bugs/welcome.shtml.
10. More information is available at http://www.nuff.ox.ac.uk/users/shephard/ox/.
11. Notice that in the unit root case, the parameters of the LMSV model are not identified.
12. The series is obtained from http://www.spglobal.com/indexmaineuro350.html.
13. It is possible that this series is too short as to find long memory.

References

Abramowitz, M. and Stegun, N. C. (1970). *Handbook of Mathematical Functions*. New York: Dover Publications Inc.

Aguilar, O. and West, M. (2000). Bayesian dynamic factor models and variance matrix discounting for portfolio allocation. *Journal of Business and Economic Statistics* 18: 338 357.

Alizadeh, S., Brandt, M. W. and Diebold, F. X. (2002). Range based estimation of stochastic volatility models. *Journal of Finance* 57: 1047 1091.

Anderson, J. (2001). On the normal inverse Gaussian stochastic volatility model. *Journal of Business and Economic Statistics* 19: 44 54.

Andersen, T. G. (1994a). Comment on E. Jacquier, N. Polson and E. Rossi. *Journal of Business and Economic Statistics* 12: 389 392.

Andersen, T. G. (1994b). Stochastic autoregressive volatility: a framework for volatility modeling. *Mathematical Finance* 4: 75 102.

Andersen, T. G. and Bollerslev, T. (1997). Heterogeneous information arrivals and return volatility dynamics: uncovering the long run in high frequency returns. *Journal of Finance* 52: 975 1005.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association* 96: 42 55.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica* 71: 579 626.

Andersen, T. G., Chung, H. and Sørensen, B. E. (1999). Efficient method of moments estimation of a stochastic volatility model: a Monte Carlo study. *Journal of Econo metrics* 91: 61 87.

Andersen, T. G. and Lund, J. (1997). Estimating continuous time stochastic volatility models of the short term interest rate. *Journal of Econometrics* 77: 343 378.

Andersen, T. G. and Sørensen, B. E. (1996). GMM Estimation of a stochastic volatility model: a Monte Carlo study. *Journal of Business and Economic Statistics* 14: 328 352.

Arteche, J. (2004). Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models. *Journal of Econometrics* 119: 131 154.

Bai, X., Russell, J. R. and Tiao, G. (2003). Kurtosis of GARCH and stochastic volatility models with non normal innovations. *Journal of Econometrics* 114: 349 360.

Bansal, R., Gallant, A. R., Hussey, R. and Tauchen, G. E. (1993). Computational aspects of nonparametric simulation estimation. In D.A. Belsley (ed.), *Computational Techniques for Econometrics and Economic Analysis*, pp. 3 22. Boston, MA: Kluwer Academic Publishers.

Bansal, R., Gallant, A. R., Hussey, R. and Tauchen, G. E. (1995). Nonparametric estima tion of structural models for high frequency marcket data. *Journal of Econometrics* 66: 251 287.

Barndorff Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics* 24: 1 13.

Barndorff Nielsen, O. E., Nicolato, E. and Shephard, N.G. (2002). Some recent develop ments in stochastic volatility modelling. *Quantitative Finance* 2: 11 23.

Barndorff Nielsen, O. E. and Shephard, N. G. (2001). Non Gaussian Ornstein Uhlenbeck based models and some of their uses in financial econometrics. *Journal of the Royal Statistical Society, Series B* 63: 167 241.

Berg, A., Meyer, R. and Yu, J. (2004). Deviance information criterion for comparing stochastic volatility models. *Journal of Business and Economic Statistics* 22: 107 120.

Black, R. (1986). Noise. *Journal of Finance* 3: 529 543.

Bollerslev, T. and Wright, J. H. (2000). Semiparametric estimation of long memory volatility dependencies: the role of high frequency data. *Journal of Econometrics* 98: 81 106.

Bollerslev, T. and Wright, J. H. (2001). High frequency data, frequency domain inference and volatility forecasting. *Review of Economics and Statistics* 83: 596 602.

Brandt, M. W. and Kang, Q. (2004). On the relationship between the conditional mean and volatility of stocks returns: a latent VAR approach. *Journal of Financial Economics* 72: 217 258.

Breidt, F. J. and Carriquiry, A. L. (1996). Improved quasi maximum likelihood estimators for stochastic volatility models. In A. Zellner and J.S.Lee (eds.), *Modelling and Prediction: Honoring Seymour Geisel*, pp. 228 247. New York: Springer Verlag.

Breidt, F. J., Crato, N. and de Lima, P. J. F. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics* 83: 325 348.

Brigo, D. and Hanzon, B. (1998). On some filtering problems arising in mathematical finance. *Insurance: Mathematics and Economics* 22: 53 64.

Calzolari, G., Fiorentini, G. and Sentana, E. (2000). Constrained EMM and indirect inference estimation. CEMFI Working Paper no. 0005.

Calzolari, G., Fiorentini, G. and Sentana, E. (2001). Indirect inference estimation of conditionally heteroskedastic factor models. CEMFI Unpublished Paper.

Capobianco, E. (1996). State space stochastic volatility models: a review of estimation algorithms. *Applied Stochastic Models and Data Analysis* 12: 265 279.

Carnero, A., Peña, D. and Ruiz, E. (2003). Persistence and curtosis in autoregressive stochastic volatility models (manuscript, Universidad Carlos III de Madrid).

Carter, C. K. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika* 81: 541 553.

Chan, N. H. and Petris, G. (2000). Long memory stochastic volatility: a Bayesian approach. *Communications in Statistics. Theory and Methods* 29: 1367 1378.

Chernov, M., Gallant, A. R., Ghysels, E. and Tauchen, G. (2003). Alternative models for stock price dynamics. *Journal of Econometrics* 116: 225 257.

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.

Chib, S., Nardari, F. and Shephard, N. (2002a). Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics* 108: 281 316.

Chib, S., Nardari, F. and Shephard, N. (2002b). Analysis of high dimensional multivariate stochastic volatility models (manuscript).

Crato, N. and Ray, B. (2002). Semi parametric smoothing estimators of long memory processes with added noise. *Journal of Statistical Planning and Inference* 105: 283 297.

Dai, Q. and Singleton, K. J. (2001). Specification analysis of affine term structure models. *Journal of Finance* 55: 1943 1978.

Danielsson, J. (1994). Stochastic volatility in asset prices: estimation with simulated maximum likelihood. *Journal of Econometrics* 64: 375 400.

Danielsson, J. (1998). Multivariate stochastic volatility models: estimation and a com parison with VGARCH models. *Journal of Empirical Finance* 5: 155 173.

Danielsson, J. and Richard, J.F. (1993). Accelerated Gaussian importance sampler with application to dynamic latent variable models. *Journal of Applied Econometrics* 8 (Suppl.): 153 173.

Deo, R. S. (2002). On GMM and QML estimation for the long memory stochastic volatility model (manuscript).

Deo, R. and Hurvich, C. (2001). On the log periodogram regression estimator of the memory parameter in the long memory stochastic volatility models. *Econometric Theory* 17: 686 710.

Ding, Z., Granger, C. W. J. and Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1: 83 106.

Duffie, D. and Singleton, K. J. (1993). Simulated moments estimation of Markov models of asset prices. *Econometrica* 61: 929 952.

Durbin, J. and Koopman, S. J. (1997). Monte Carlo maximum likelihood estimation for non Gaussian state space models. *Biometrika* 84: 669 684.

Durbin, J. and Koopman, S. J. (2000). Time series analysis of non Gaussian observations based on state space models from both classical and Bayesian perspectives (with discussion). *Journal of the Royal Statistical Society B* 62: 3 56.

Engle, R. F. (1994). Indirect inference on volatility diffusions and stochastic volatility models (manuscript, University of California at San Diego).

Engle, R. F. and Lee, G. (1996). Estimating diffusions models of stochastic volatility. In: P.E. Rossi (ed.), *Modelling Stock Market Volatility: Bridging the Gap to Continuous Time*, pp. 333 355. San Diego: Academic Press.

Eraker, B., Johannes, M. and Polson, N. (2003). The impact of jumps in volatility and returns. *Journal of Finance* 58: 1269 1300.

Fiorentini, G., León, A. and Rubio, G. (2002). Estimation and empirical performance of Heston's stochastic volatility model: The case of a thinly traded market. *Journal of Empirical Finance* 9: 225 255.

Fleming, J., Kirby, C. and Ostdiek, B. (1998). Information and volatility linkages in the stock, bond and money markets. *Journal of Financial Economics* 49: 111 137.

Fridman, M. and Harris, L. (1998). A maximum likelihood approach for non Gaussian stochastic volatility models. *Journal of Business and Economic Statistics* 16: 284 291.

Fuller, W. A. (1996). *Introduction to Statistical Time Series*. New York: Wiley.

Gallant, A. R., Hsieh, D. and Tauchen, G. (1997). Estimation of stochastic volatility model with diagnostics. *Journal of Econometrics* 81: 159 192.

Gallant, A. R. and Long, J. R. (1997). Estimating stochastic differential equations efficiently by minimum Chi square. *Biometrika* 84: 125 141.

Gallant, A. R. and Tauchen, G. (1996). Which moments to match? *Econometric Theory* 12: 657 681.

Gallant, A. R. and Tauchen, G. (1998). Reprojecting partially observed systems with application to interest rate diffusions. *Journal of the American Statistical Association* 93: 10 24.

Geweke, J. (1994). Bayesian comparison of econometric models. Working Paper, Federal Reserve Bank of Minneapolis, Minnesota.

Geweke, J. and Porter Hudak, S. (1983). Estimation and application of long memory time series models. *Journal of Time Series Analysis* 4: 221 238.

Ghysels, E., Harvey, A. C. and Renault, E. (1996). Stochastic volatility. In G.S. Maddala and C. R. Rao (eds.), *Statistical Methods in Finance*, pp. 119 191. North Holland, Amsterdam.

Ghysels, E. and Jasiak, J. (1996). Stochastic volatility and time deformation: an application to trading volume and leverage effects (manuscript, Université de Montréal).

Gourieroux, C., Monfort, A. and Renault, E. (1993). Indirect inference. *Journal of Applied Econometrics* 8: 85 118.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* 50: 1029 1054.

Harvey, A. C. (1998). Long memory in stochastic volatility. In J. Knight and S. Satchell (eds.), *Forecasting Volatility in Financial Markets*, pp. 307 320. Oxford: Butterworth Heinemann.

Harvey, A. C., Ruiz, E. and Shephard, N. G. (1994). Multivariate stochastic variance models. *Review of Economic Studies* 61: 247 264.

Harvey, A. C. and Shephard, N. G. (1993). The econometrics of stochastic volatility. London School of Economics, Working Paper 166.

Harvey, A. C. and Shephard, N. G. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business and Economic Statistics* 14: 429 434.

Harvey, A. C. and Streibel, M. (1998). Testing for a slowly changing level with special reference to stochastic volatility. *Journal of Econometrics* 87: 167 189.

Hsu, N. J. and Breidt, F. J. (1997). Bayesian analysis of long memory stochastic volatility models. Department of Statistics, Iowa State University, Technical Report, 97 94.

Hwang, S. and Satchell, S. E. (2000). Market risk and the concept of fundamental volatility: measuring volatility across asset and derivative markets and testing for the impact of derivatives markets on financial markets. *Journal of Banking and Finance* 24: 759 785.

Islas Camargo, A. and Venegas Martínez, F. (2003). Long memory volatility in Latin American stock markets (manuscript, Department of Statistics, ITAM).

Jacquier, E., Polson, N. G. and Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics* 12: 371 417.

Jacquier, E., Polson, N. G. and Rossi, P. E. (1995). Models and prior distributions for multivariate stochastic volatility. Working Paper no. 95s 18, CIRANO.

Jacquier, E., Polson, N. G. and Rossi, P. E. (2004). Bayesian analysis of stochastic volatility models with fat tails and correlated errors. *Journal of Econometrics* 122: 185 212.

Jensen, M. J. (2003). Bayesian inference of long memory stochastic volatility via wavelets (manuscript, University of Missouri).

Jiang, G. J. and van der Sluis, P. J. (1998). Forecasting volatility under multivariate stochastic volatility model via reprojections (manuscript, University of Amsterdam).

Jiang, G. J. and van der Sluis, P. J. (2000). Index option pricing models with stochastic volatility and stochastic interest rates. *European Finance Review* 3: 273 310.

Kim, S. and Shephard, N. (1994). Stochastic volatility. *Optimal Likelihood Inference and Comparison with ARCH Models* (Unpublished Paper, Nuffield College, Oxford).

Kim, S., Shephard, N. and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies* 45: 361 393.

Kitagawa, G. (1987). Non Gaussian state space modeling of nonstationary time series. *Journal of the American Statistical Association* 82: 1032 1063.

Knight, J. L. and J. Yu (2002). The empirical characteristic function in time series estimation. *Econometric Theory* 18: 691 721.

Knight, J. L., Satchell, S. E. and J. Yu (2002). Estimation of the stochastic volatility model by the empirical characteristic function method. *Australian and New Zealand Journal of Statistics* 44: 319 335.

Koopman, S. J., Shephard, N. and Doornik, J. (1999). Statistical algorithms for models in state space using SsfPack 2.2. *Econometrics Journal* 2: 113 166.

Koopman, S. J. and Uspensky, E. H. (2002). The stochastic volatility in mean model: empirical evidence from international stock markets. *Journal of Applied Econometrics* 17: 667 689.

Lien, D. and Wilson, B. K. (2001). Multiperiod hedging in the presence of stochastic volatility. *International Review of Financial Analysis* 10: 395 406.

Liesenfeld, R. (1998). Dynamic bivariate mixture models: modeling the behavior of prices and trading volume. *Journal of Business and Economic Statistics* 16: 101 109.

Liesenfeld, R. (2001). A generalized bivariate mixture model for stock price volatility and trading volume. *Journal of Econometrics* 104: 141 178.

Liesenfeld, R. and Jung, R. C. (2000). Stochastic volatility models: conditional normality versus heavy tailed distributions. *Journal of Applied Econometrics* 15: 137 160.

Liesenfeld, R. and Richard, J. F. (2003). Univariate and multivariate stochastic volatility models: estimation and diagnostics. *Journal of Empirical Finance* 207: 1 27.

Lopes, H. F. and Migon, H. S. (2003). Comovements and contagion in emergent markets: stock indexes volatilities (manuscript, Universidade Federal do Rio de Janeiro).

Mahieu, R. and Bauer, R. (1998). A Bayesian analysis of stock volatility and trading volume. *Applied Financial Economics* 8: 671 687.

Mahieu, R. J. and Schotman, P. C. (1998). An empirical application of stochastic volatility models. *Journal of Applied Econometrics* 13: 333 360.

McMillan, D. G. (2001). Common stochastic volatility trend in European exchange rates. *Applied Economics Letters* 8: 605 608.

Melino, A. and Turnbull, S. M. (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics* 45: 239 265.

Meyer, R. and Yu, J. (2000). BUGS for a Bayesian analysis of stochastic volatility models. *Econometrics Journal* 3: 198 215.

Monfardini, C. (1998). Estimating stochastic volatility models through indirect inference. *Econometrics Journal* 1: 113 128.

Nelson, D. B. (1988). The time series behavior of stock market volatility and returns. (Unpublished Ph D Dissertation, Massachusetts Institute of Technology, Cambridge, MA).

Pastorello, S., Renault, E. and Touzi, N. (1994). Statistical inference for random variance option pricing. CREST Dipartment de la Recherce INSEE, DP.

Pérez, A. (2000). Estimación e Identificación de Modelos de Volatilidad Estocástica con Memoria Larga. PhD Thesis, Departamento de Economía Aplicada (Estadística y Econometría), Universidad de Valladolid. UMI Dissertation Services, ISBN 0 493 58794 2.

Pérez, A. and Ruiz, E. (2001). Finite sample properties of a QML estimator of stochastic volatility models with long memory. *Economics Letters* 70: 157 164.

Pérez, A. and Ruiz, E. (2003). Properties of the sample autocorrelations of non linear transformations in long memory stochastic volatility models (manuscript, Universidad Carlos III de Madrid).

Pitt, M. K. and Shephard, N. (1999a). Filtering via simulation: auxiliary particle filter. *Journal of the American Statistical Association* 94: 590 599.

Pitt, M. K. and Shephard, N. (1999b). Time varying covariances: a factor stochastic volatility approach (with discussion). In J. Bernardo, J.O. Berger, A.P. Dawid and A. F. M. Smith (eds.), *Bayesian Statistics*, 6, pp. 547 570. Oxford: Oxford University Press.

Ray, B. K. and Tsay, R. S. (2000). Long range dependence in daily stock volatilities. *Journal of Business and Economic Statistics* 18: 254 262.

Robinson, P. M. (2001). The memory of stochastic volatility models. *Journal of Econometrics* 101: 195 218.

Ruiz, E. (1994). Quasi maximum likelihood estimation of stochastic volatility models. *Journal of Econometrics* 63: 289 306.

Sandmann, G. and Koopman, S. J. (1998). Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics* 87: 271 301.

Shephard, N. G. (1993). Fitting non linear time series models, with application to stochastic variance models. *Journal of Applied Econometrics* 8: 135 152.

Shephard, N. (1994). Partial non Gaussian state space. *Biometrika* 81: 115 131.

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In O.E. Barndorff Nielsen, D. R. Cox and D. V.Hinkley (eds.), *Statistical Models in Econometrisc, Finance and Other Fields*, pp. 1 67. London: Chapman & Hall.

Shephard, N. G. (2000). Discussion of Durbin and Koopman (2000). *Journal of the Royal Statistical Society, Series B* 62: 30 32.

Shephard, N. G. and Kim, S. (1994). Comment on E. Jacquier, N. Polson and E. Rossi (1994). *Journal of Business and Economic Statistics* 12: 406 410.

Shephard, N. and Pitt, M. K. (1997). Likelihood analysis of non Gaussian parameter driven models. *Biometrica* 84: 653 667.

Singleton, K. J. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics* 102: 111 141.

van der Sluis, P. J. (1997). Emmpack 1.01: C/C++ code for use with Ox for estimation of univariate stochastic volatility models with the efficient method of moments. *Studies in Nonlinear Dynamics and Econometrics* 2.3: 77 94.

So, M. K. P. (1999). Time series with additive noise. *Biometrika* 86: 474 482.

So, M. K. P. (2002). Bayesian analysis of long memory stochastic volatility models. *Sankhya. The Indian Journal of Statistics* 64(B) 1: 1 10.

So, M. K. P., Lam, K. and Li, W.K. (1998). A stochastic volatility model with Markov switching. *Journal of Business and Economic Statistics* 16: 244 253.

So, M. K. P. and Li, W.K. (1999). Bayesian unit root testing in stochastic volatility models. *Journal of Business and Economic Statistics* 17: 491 496.

So, M. K. P., Li, W.K. and Lam, K. (2002). A threshold stochastic volatility model. *Journal of Forecasting* 21: 473 500.

Steel, M. F. J. (1998). Bayesian analysis of stochastic volatility models with flexible tails. *Econometric Reviews* 17 (2): 109 143.

Tauchen, G. (1997). New minimum Chi square methods in empirical finance. In D.M. Kreps and K. F. Wallis (eds.), *Advances in Economics and Econometics: Theory and Applications, Seventh World Congress*, pp. 279 317. Cambridge: Cambridge University Press.

Taylor, S. (1986). *Modeling Financial Time Series*. Chichester: Wiley.

Taylor, S. (1994). Modeling stochastic volatility: a review and comparative study. *Mathematical Finance* 4: 183 204.

Tierney, L. (1994). Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics* 22: 1701 1762.

Tims, B. and Mahieu, R. (2003). A range based model for exchange rate volatility. WP ERS 2003 002 F&A, ERIM.

Vetzal, K. V. (1997). Stochastic volatility, movements in short term interest rates, and bond options values. *Journal of Banking and Finance* 21: 169 196.

Watanabe, T. (1999). A non linear filtering approach to stochastic volatility models with an application to daily stock returns. *Journal of Applied Econometrics* 14: 101 121.

Watanabe, T. and Asai, M. (2001). Stochastic volatility models with heavy tailed distributions: A Bayesian approach. IMES Discussion Paper Series 2001 E 17.

Wong, C. K. J. (2002). Latent factor models of high frequency financial time series. Unpublished PhD Thesis. University of Oxford.

Wright, J. H. (1999). A new estimator of the fractionally integrated stochastic volatility model. *Economics Letters* 63: 295 303.

Yu, J. (2002a). Forecasting volatility in the New Zealand stock market. *Applied Financial Economics* 12: 193 202.

Yu, J. (2002b). MCMC methods for estimating stochastic volatility models with leverage effects: comments on Jacquier, Polson and Rossi (2002) (manuscript, University of Auckland).

Yu, J., Yang, Z. and Zhang, X. (2002). A class of nonlinear stochastic volatility models and its implications on pricing currency options (manuscript, University of Auckland).