Quasi-maximum likelihood estimation of stochastic volatility models

Esther Ruiz

Departamento de Estadística y Econometría, Universidad Carlos III, C/Madrid 126-128, 28903 Getafe, Spain

Abstract

Changes in variance or volatility over time can be modelled using stochastic volatility (SV) models. This approach is based on treating the volatility as an unobserved variable, the logarithm of which is modelled as a linear stochastic process, usually an auto-regression. This article analyses the asymptotic and finite sample properties of a Quasi-Maximum Likelihood (QML) estimator based on the Kalman filter. The relative efficiency of the QML estimator when compared with estimators based on the Generalized Method of Moments is shown to be quite high for parameter values often found in empirical applications. The QML estimator can still be employed when the SV model is generalized to allow for distributions with heavier tails than the normal. SV models are finally fitted to daily observations on the yen/dollar exchange rate.

Key words: Exchange rates; Generalized method of moments; Kalman filter; Quasi-maximum likelihood; Stochastic volatility
JEL classification: C22

1

1. Introduction

Time series models with heteroscedastic errors are receiving an increasing attention in the econometric literature. These models are specially useful for modelling high-frequency financial time series, such as stock returns and exchange rates. In the simplest set up, the series of interest is a white noise process, with unit variance, multiplied by a factor σ_t, known in the financial literature as *volatility*. That is,

$$y_t = \sigma_t \varepsilon_t. \tag{1}$$

There are different ways for modelling changes in volatility over time. Some of the most popular models in the literature are those based on the Autoregressive Conditional Heteroscedasticity (ARCH), inspired by Engle (1982); see Bollerslev et al. (1992) for a detailed review of these models. All ARCH-based models share the property that the distribution of y_t conditional on past values of the series, up to and including time $t - 1$, is Gaussian with variance σ_t^2. Consequently, the volatility is observable at time $t - 1$.

Alternatively, the volatility may be treated as an unobserved variable, the logarithm of which is modelled as a linear stochastic process, such as an autoregression. Models of this kind are called stochastic volatility (SV) models; see Taylor (1991) for an excellent review on SV models.

ARCH-type models and SV models have similar statistical properties. However, they are different with respect to the observability of σ_t^2 at time $t - 1$; see Andersen (1992) for a detailed discussion on this subject.

SV models fit quite naturally into the theoretical framework within which much of the modern finance theory, in relation to option valuation, has been developed; see, for example, Chesney and Scott (1989). Another feature of these models is that they can be naturally generalized to multivariate series as in Harvey et al. (1992). However, they are not conditionally Gaussian and, therefore, their estimation and statistical handling may present some difficulties. The estimation of SV models has usually been carried out by variants of the method of moments. The aim of this article is to analyse the properties of a Quasi-Maximum Likelihood (QML) estimator of SV models based on the Kalman filter and proposed independently by Nelson (1988) and Harvey et al. (1992) and compare them with the properties of some estimators based on the method of moments.

The article is organised as follows. Section 2 sets out some of the basic ideas of univariate SV models. In Section 3, the QML estimator is described and its asymptotic and finite-sample properties are found. In Section 4, the asymptotic and finite-sample properties of two estimators based on the method of moments are analysed. It is shown that for parameter values likely to arise in practice when analysing very-high-frequency financial time series, these estimators are

less efficient than the QML estimator. Section 5 presents an empirical application using a series of daily yen/dollar exchange rates. The conclusions are given in Section 6.

2. Stochastic volatility processes

A simple stationary SV model is given by

$$y_t = \varepsilon_t \exp\{h_t/2\}, \tag{2a}$$

$$h_t = \gamma + \phi h_{t-1} + \eta_t, \qquad \eta_t \sim \text{NID}(0, \sigma_\eta^2), \tag{2b}$$

where $h_t = \ln(\sigma_t^2)$, ε_t is a white noise process with unit variance, generated independently of η_t and $|\phi| < 1$. Working with logarithms ensures that σ_t^2 is always positive.

Transforming y_t by taking logarithms of the squares, we obtain a linear state space model

$$\ln(y_t^2) = \text{E}(\ln(\varepsilon_t^2)) + h_t + \xi_t, \tag{3a}$$

$$h_t = \gamma + \phi h_{t-1} + \eta_t, \tag{3b}$$

where $\xi_t = \ln(\varepsilon_t^2) - \text{E}(\ln(\varepsilon_t^2))$. ξ_t is a non-Gaussian, zero mean, white noise, and its statistical properties depend on the distribution of ε_t. From (3) it is possible to observe that, $\ln(y_t^2)$ is equivalent to an $\text{ARMA}(1,1)$ process with a non-Gaussian noise.

If $\varepsilon_t \sim \text{NID}(0,1)$, as it is often assumed in the literature, the mean and variance of $\ln(\varepsilon_t^2)$ are known to be $\psi(\frac{1}{2}) - \ln(\frac{1}{2}) \approx -1.27$ and $\pi^2/2$ respectively, where $\psi(\cdot)$ is the Digamma function; see Abramovitz and Stegun (1970).

Some authors also assume that ε_t have a Student-t distribution; see Harvey et al. (1992). Let ε_t be a t-variable written as

$$\varepsilon_t = \zeta_t/\kappa_t^{1/2}, \tag{4}$$

where ζ_t is a standard normal variate and $\nu\kappa_t$ is distributed, independently of ζ_t, as a χ^2 with ν degrees of freedom. Then $\ln(\varepsilon_t^2) = \ln(\zeta_t^2) - \ln(\kappa_t)$. It follows from results in Abramovitz and Stegun (1970) that, in this case, $\text{E}[\ln(\varepsilon_t^2)] \approx -1.27 - \psi(\nu/2) + \ln(\nu/2)$ and $\text{var}[\ln(\varepsilon_t^2)] = \pi^2/2 + \psi'(\nu/2)$, where $\psi'(\cdot)$ is the Trigamma function.

The SV model in (2) can be generalised so that h_t follows any stationary ARMA process. Alternatively, h_t can be allowed to follow a random walk. The corresponding SV model is then given by

$$y_t = \varepsilon_t \exp\{h_t/2\}, \tag{5a}$$

$$h_t = h_{t-1} + \eta_t. \tag{5b}$$

3

In this case, $\ln(y_t^2)$ is a random walk plus noise, and the best linear predictor of the current value of h_t is an exponential weighted moving average (EWMA) of past values of $\ln(y_t^2)$.

3. Quasi-maximum likelihood estimation

The QML approach for estimating SV models has been proposed independently by Nelson (1988) and Harvey et al. (1992) and is based on the Kalman filter. This is applied to $\ln(y_t^2)$ to obtain one-step-ahead errors and their variances. These are then used to construct a quasi-likelihood function. Because $\ln(y_t^2)$ is not Gaussian, the Kalman filter yields minimum mean square linear estimators (MMSLE) of h_t and future observations rather than minimum mean square estimators (MMSE).

3.1. Stationary case

Consider the SV model (2) with $\varepsilon_t \sim \text{NID}(0, 1)$. In order to estimate the parameters γ, ϕ, and σ_η^2, consider the linear state space model in (3), which in this case is given by

$$\ln(y_t^2) = -1.27 + h_t + \xi_t, \tag{6a}$$

$$h_t = \gamma + \phi h_{t-1} + \eta_t. \tag{6b}$$

where $\sigma_\xi^2 = \pi^2/2$.

Estimation of model (6) can be carried out by QML by treating ξ_t as though it were $\text{NID}(0, \pi^2/2)$. The Kalman filter may then be used to obtain the prediction error decomposition form of the Gaussian likelihood, which has to be numerically maximised. Notice that model (6) can be written

$$y_t^* = h_t + \xi_t, \tag{7a}$$

$$h_t = \phi h_{t-1} + \eta_t, \tag{7b}$$

where y_t^* is equal to $\ln(y_t^2)$ corrected by its mean given by $\gamma^* = -1.27 + \gamma/(1 - \phi)$.

γ^* can be consistently estimated by the sample mean of $\ln(y_t^2)$. Moreover, the sample mean is a QML estimator of γ^* uncorrelated with the QML estimator of the stochastic part of the model; see Harvey (1989). In what follows, we treat γ^*, or equivalently γ, as known and concentrate on the estimation of $\Psi = (\phi, \sigma_\eta^2)'$.

The standard theory for the estimation of unobserved component time series models with nonnormal errors applies to the estimates of Ψ. Using the results in Dunsmuir (1979) and since η_t and ξ_t are martingale differences and have finite

4

moments up to at least order four, the asymptotic distribution of the QML estimator of Ψ, denoted by $\hat{\Psi}$, is given by

$$\sqrt{T}(\hat{\Psi} - \Psi) \xrightarrow{\text{d}} N(0, C(\Psi)),$$

where T is the sample size and the analytic expression for $C(\Psi)$ is given in the appendix.

Table 1 shows the finite-sample approximation to the asymptotic standard errors (ASEs), given by $\sqrt{C(\Psi)}/\sqrt{T}$, for different parameter values and sample sizes. The finite-sample properties of $\hat{\Psi}$ have been studied by carrying out several Monte Carlo experiments. Table 2 shows the results of some of these experiments. The small-sample bias in the QML estimates for ϕ is generally of relatively minor order in the present context with a relatively large number of observations. The QML estimates for σ_η^2 are also slightly biased in finite samples. Both biases tend to increase when ϕ decreases or σ_η^2 decreases. Comparison between the ASEs in Table 1 and the Monte Carlo standard errors shows that for $T = 500$ the ASEs of $\hat{\phi}$ and $\hat{\sigma}_\eta^2$ underestimate the observed standard errors by quite a large amount when $\phi = 0.9$ and $\sigma_\eta^2 = 0.09$. This may be due to the fact that both parameters are getting close to the boundary of the permissible parameter space. The ASEs of $\hat{\sigma}_\eta^2$ are, in general, a better approximation to the observed standard errors.

3.2. Random walk plus noise case

When h_t follows a random walk as in model (5), the Kalman filter approach is still valid if the restriction $\phi = 1$ is imposed. The only difference is that the first

Table 1

Asymptotic standard deviations of QML estimator of stationary stochastic variance model

	$T = 500$	$T = 3000$	$T = 6000$
$\phi = 0.9$	0.0306	0.0125	0.0088
$\sigma_\eta^2 = 1$	0.2890	0.1180	0.0834
$\phi = 0.7$	0.1033	0.0422	0.0298
$\sigma_\eta^2 = 1$	0.4802	0.1960	0.1386
$\phi = 0.9$	0.0792	0.0323	0.0229
$\sigma_\eta^2 = 0.09$	0.1004	0.0410	0.0290
$\phi = 0.7$	0.6761	0.2760	0.1951
$\sigma_\eta^2 = 0.09$	0.3407	0.1391	0.0983

Table 2

Monte Carlo experiments on QML estimator of the stationary stochastic variance model: Average value with standard deviations in parentheses (1000 replications)

	$T = 500$	$T = 3000$	$T = 6000$
$\phi = 0.9$	0.8927	0.9004	0.9010
	(0.0342)	(0.0170)	(0.0150)
$\sigma_\eta^2 = 1$	1.0276	0.9942	0.9964
	(0.3117)	(0.1207)	(0.0913)
$\phi = 0.7$	0.6840	0.6982	0.6979
	(0.1081)	(0.0434)	(0.0230)
$\sigma_\eta^2 = 1$	1.0612	1.0030	1.0076
	(0.5015)	(0.1935)	(0.1400)
$\phi = 0.9$	0.7464	0.8932	0.8968
	(0.2691)	(0.0406)	(0.0254)
$\sigma_\eta^2 = 0.09$	0.2178	0.0996	0.0949
	(0.3314)	(0.0528)	(0.0329)
$\phi = 0.7$	0.3688	0.5362	0.6138
	(0.2857)	(0.2841)	(0.2314)
$\sigma_\eta^2 = 0.09$	0.2583	0.1481	0.1305
	(0.3627)	(0.1532)	(0.1184)

observation is used to initialise the Kalman filter, whereas when $\phi < 1$, the unconditional distribution of h_t is available at $t = 0$.

Assuming that $\varepsilon_t \sim \text{NID}(0, 1)$, consider the following state space model:

$$\ln(y_t^2) = -1.27 + h_t + \xi_t, \tag{8a}$$

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

where $\sigma_\xi^2 = \pi^2/2$. In this case, the model contains only one unknown parameter, σ_η^2. Denote by $\hat{\sigma}_\eta^2$ the QML estimator of σ_η^2. The asymptotic distribution of $\sqrt{T}(\hat{\sigma}_\eta^2 - \sigma_\eta^2)$ is normal with zero mean and asymptotic variance $C(\sigma_\eta^2)$. The analytic expression for $C(\sigma_\eta^2)$ appears in the Appendix.

Table 3 shows the approximation in finite-sample sizes to the ASEs of $\hat{\sigma}_\eta^2$, for different values of σ_η^2 and different sample sizes.

The results of several Monte Carlo experiments, carried out to analyse the finite-sample properties of $\hat{\sigma}_\eta^2$, appear in Table 4. The QML estimates of σ_η^2 show a slight finite-sample bias which tends to increase when σ_η^2 decreases. Moreover, comparing Tables 3 and 4, one observes that, for small values of σ_η^2, the ASEs underestimate the observed standard errors. The problem may be that, when σ_η^2 is close to zero, the distribution of $\hat{\sigma}_\eta^2$ may be some way from normality; see

Table 3
Standard deviation of QML estimator of random walk volatility model

σ_η^2	$T = 500$	$T = 3000$	$T = 6000$
0.09	0.03196	0.01305	0.00922
0.01	0.00603	0.00246	0.00174
0.0009	0.00098	0.00040	0.00028
0.0001	0.00019	0.00008	0.00005

Table 4
Monte Carlo experiments on QML estimator of the random walk volatility model: Average value with standard deviations in parentheses (1000 replications)

σ_η^2	$T = 500$	$T = 3000$	$T = 6000$
0.09	0.09028	0.08949	0.09005
	(0.03356)	(0.01269)	(0.00903)
0.01	0.01015	0.00996	0.01002
	(0.00689)	(0.00247)	(0.00175)
0.0009	0.00101	0.00089	0.00091
	(0.00148)	(0.00042)	(0.00028)
0.0001	0.00028	0.00010	0.00010
	(0.00083)	(0.00009)	(0.00006)

Harvey (1989). For relatively large values of σ_η^2, the ASEs are good approximations to the finite-sample standard deviations.

3.3. Heavy-tailed distributions

When ε_t is a t-variable as given by (4) and h_t is stationary, consider the state space model in (7) with $\gamma^* = -1.27 - \psi(v/2) + \log(v/2) - \gamma/(1 - \phi)$ and $\sigma_\xi^2 = \pi^2/2 + \psi'(v/2)$. Then, the model can be estimated by estimating σ_ξ^2, γ^*, ϕ, and σ_η^2 with the restriction $\sigma_\xi^2 \geqslant \pi^2/2$. The estimated variance of ξ_t implies a value of v. As in the Gaussian case, γ^* can be estimated by the sample mean and therefore we concentrate on the estimation of $\Psi^* = (\sigma_\xi^2, \phi, \sigma_\eta^2)'$. Estimates of Ψ^* can be obtained by treating ξ_t as though it were $\mathrm{NID}(0, \sigma_\xi^2)$ and maximising the resulting quasi-likelihood function. In this case, the asymptotic distribution of the QML estimator of Ψ^* is given by

$$\sqrt{T}(\hat{\Psi}^* - \Psi^*) \xrightarrow{\mathrm{d}} \mathrm{N}(0, C(\Psi^*)).$$

The elements of $C(\Psi^*)$ appear in the Appendix.

It is also possible to estimate SV models by assuming that v is known. In this case, there is no need to estimate σ_ξ^2. Surprisingly, for the typical parameter values often found in empirical applications, the estimates of ϕ and σ_η^2 are less efficient in this case. Table 5 reports values of ASEs of the QML estimators of $\Psi = (\phi, \sigma_\eta^2)'$ obtained by assuming that v is known and when this parameter is estimated. It can be observed that the loss in efficiency is greater for $\hat{\sigma}_\eta^2$ than for $\hat{\phi}$. For the parameter values considered in Table 5, the loss in efficiency for $\hat{\phi}$ can be shown to be quite small. However, for $\hat{\sigma}_\eta^2$ the loss could be as big as 3% in some cases. This result seems counterintuitive at first sight. However, since the proposed estimators are QML and not maximum likelihood estimators, they are not fully efficient. Consequently, it may be possible that the QML estimator of σ_η^2 when both variances are estimated is more efficient than when σ_ξ^2 is known.

With respect to hypothesis testing, an interesting hypothesis to test is whether ε_t is normal, i.e., H_0: $\sigma_\xi^2 = \pi^2/2$. One possible test statistic could be a quasi-Likelihood Ratio (LR) test. Since under the null hypothesis, σ_ξ^2 is on the

Table 5

Asymptotic standard deviations with Student-t disturbances and relative efficiency computed as the ratio of standard deviations

	Estimating v	v known	Efficiency loss
(a) Stationarity			
$v = 6$			
$\phi = 0.95$	0.6634	0.6708	1.0112
$\sigma_\eta^2 = 0.09$	1.3577	1.3847	1.0199
$v = 10$			
$\phi = 0.95$	0.6556	0.6647	1.0139
$\sigma_\eta^2 = 0.09$	1.3342	1.3670	1.0246
$v = 6$			
$\phi = 0.95$	2.2472	2.2785	1.0139
$\sigma_\eta^2 = 0.01$	0.6645	0.6792	1.0221
$v = 6$			
$\phi = 0.99$	0.1646	0.1650	1.0024
$\sigma_\eta^2 = 0.09$	0.8275	0.8367	1.0111
(b) Random walk			
$v = 6$			
$\sigma_\eta^2 = 0.09$	0.7183	0.7384	1.0280
$v = 6$			
$\sigma_\eta^2 = 0.01$	0.1367	0.1379	1.0088
$v = 6$			
$\sigma_\eta^2 = 0.001$	0.0242	0.0243	1.0041

boundary of the admissible parameter space, the distribution of the LR test is given by

$$LR \sim \tfrac{1}{2}\chi_0^2 + \tfrac{1}{2}\chi_1^2.$$

where χ_0^2 is a degenerate distribution with all its mass at the origin; see Harvey (1989). The size of the LR test can therefore be set appropriately simply by using the 2α, rather than the α, significance point of a χ_1^2 distribution for a test of size α. For example, for $\alpha = 5\%$, the corresponding critical value for the quasi-LR test statistic is 2.71. This is the value found by Bollerslev (1987) by carrying out Monte Carlo experiments.

Alternatively, by estimating σ_ξ^2 as an unrestricted parameter (i.e., without imposing the lower bound $\pi^2/2$), it is possible to allow ε_t to have a distribution with kurtosis which could be smaller or bigger than the normal kurtosis. For example, ε_t may have a generalized error distribution (GED); see Nelson (1988). The reason why the GED family of distributions could be an attractive alternative is because it includes the normal as an special case, and also includes distributions with thinner and fatter tails than the normal.

Finally, when ε_t is a t-variate and h_t follows a random walk as in (5), the corresponding SV model can be estimated by estimating $\Psi^+ = (\sigma_\xi^2 \ \sigma_\eta^2)'$ in the appropriate linear state space model. The asymptotic distribution of $\sqrt{T}(\hat{\Psi}^+ - \Psi^+)$ is also normal with zero mean and variances and covariances matrix, $C(\Psi^+)$, given in the Appendix.

4. Estimators based on the method of moments principle

The estimation of SV models has usually been carried out by variants of the method of moments. These methods have the difficulty that their efficiency depends on the choice of moments. In this section, we compare the asymptotic properties of the QML estimator of SV models with two estimators based on the method of moments principle. In what follows we will assume that $\varepsilon_t \sim NID(0, 1)$.

4.1. Estimator based on the sample variance

Consider the SV model in (5) with h_t following a random walk. The corresponding linear state space form is given by (8).

The stationary form of $y_t^+ = \ln(y_t^2)$ is given by $\Delta y_t^+ = \eta_t + \Delta \xi_t$.

Given that η_t and ξ_t are mutually uncorrelated and $\sigma_\xi^2 = \pi^2/2$, the variance of η_t is given by

$$\sigma_\eta^2 = \sigma_{\Delta y}^2 \dagger - \pi^2.$$

9

A method of moments estimator of σ_η^2 is then given by

$$\tilde{\sigma}_\eta^2 = \tilde{\sigma}_{\Delta y}^2 \dagger - \pi^2, \tag{9}$$

where $\tilde{\sigma}_{\Delta y}^2 \dagger$ is the sample variance of $\Delta\ln(y_t^2)$.

If $\sigma_\eta^2 > 0$, then $\sqrt{T}(\tilde{\sigma}_\eta^2 - \sigma_\eta^2)$ has an asymptotic normal distribution with zero mean and variance $C_2 = 2[(\sigma_\eta^2 + \pi^2)^2 + \pi^4]$.

Table 6 shows the ASEs of $\hat{\sigma}_\eta^2$ and $\sqrt{C_2}/\sqrt{T}$, together with the square root of the relative efficiency of $\tilde{\sigma}_\eta^2$ with respect to $\hat{\sigma}_\eta^2$. It is possible to observe that, for the parameter values considered, the efficiency of the method of moments estimator compared with the QML estimator is exceptionally low.

4.2. Generalized method of moments estimator

Melino and Turnbull (1990) propose an estimator of SV models based on the Generalized Method of Moments (GMM) principle. Consider the stationary SV model in (2) with $\varepsilon_t \sim \text{NID}(0, 1)$. The functions used by Melino and Turnbull to compute the GMM estimator of model (2) are given by

$$y_t^2 - E(y_t^2), \tag{10a}$$

$$y_t^4 - E(y_t^4), \tag{10b}$$

$$|y_t| - E(|y_t|), \tag{10c}$$

$$|y_t^3| - E(|y_t^3|), \tag{10d}$$

$$y_t^2 y_{t-\tau}^2 - E(y_t^2 y_{t-\tau}^2), \qquad \tau = 1, \dots, 10, \tag{10e}$$

$$|y_t y_{t-\tau}| - E|y_t y_{t-\tau}|, \qquad \tau = 1, \dots, 10. \tag{10f}$$

Denote by $\tilde{\Psi}$ the GMM estimator of Ψ. Under suitable conditions, Hansen (1982) shows that $\sqrt{T}(\tilde{\Psi} - \Psi)$ is asymptotically distributed as a normal variable with zero mean and covariance matrix given by $W = (D'V^{-1}D)^{-1}$, where $D' = \text{plim}(\partial g(\Psi)/\partial\Psi)$, $g(\Psi) = (1/T)\Sigma g_t$, g_t denote a vector whose components are functions of y_t drawn from (10), and $V = E(g_t g_t')$.

Table 6
Asymptotic standard deviation of QML and method of moments (MM) estimators of random walk volatility model

σ_η^2	QML	MM	(Rel. eff.)$^{1/2}$
0.09	0.7763	19.8294	0.0391
0.01	0.1391	19.7492	0.0070
0.0009	0.0222	19.7401	0.0011
0.0001	0.0042	19.7393	0.0002

The derivatives needed to compute D have been derived by Vetzal (1991). The moments involved in the construction of V can be derived using the moments of y_t given in an Appendix to Melino and Turnbull (1990), kindly supplied to us by the authors.

Table 7 shows the ASEs of $\tilde{\Psi}$ together with the ASEs of the QML estimator. Table 7 also shows the square root of the relative efficiency of the GMM estimator with respect to the QML estimator. Comparing the ASEs of both estimators, it is possible to observe that, in terms of efficiency, the QML estimation method performs better when ϕ is close to one and σ_η^2 is relatively big. On the other hand, when σ_η^2 and ϕ are small, the GMM estimator is more efficient. This result could be expected, because when σ_η^2 and ϕ are small, the variance of the log-volatility process, given by $\sigma_\eta^2/(1 - \phi^2)$, is very small in relation to the variance of ξ_t. Therefore, the transformation $\ln(y_t^2)$ is dominated by ξ_t, and the approximation to normality used by the Kalman filter is very poor. Also, in these circumstances, the process y_t is approximately normal, with the excess kurtosis being small. As a consequence, the sample moments are reliable estimators of the population moments, and the GMM estimator performs quite well.

However, in most empirical applications with very-high-frequency financial time series (hourly or daily), it has been observed that the parameter ϕ is very close or exactly one. The estimated values for σ_η^2 are usually between 0.01 and 2.77; see Taylor (1991). With this range of parameter values, there is little doubt

Table 7
Asymptotic standard deviations of QML and GMM estimators

	QML	GMM	(Rel. eff.)$^{1/2}$
$\phi = 0.9$	0.68	12203.65	0.00
$\sigma_\eta^2 = 1$	6.46	28903.37	0.00
$\phi = 0.7$	2.31	29.17	0.08
$\sigma_\eta^2 = 1$	10.74	20.13	0.53
$\phi = 0.9$	1.77	3.77	0.47
$\sigma_\eta^2 = 0.09$	2.24	0.86	2.60
$\phi = 0.7$	15.12	13.94	1.08
$\sigma_\eta^2 = 0.09$	7.62	1.02	7.47
$\phi = 0.95$	0.92	3.00	0.31
$\sigma_\eta^2 = 0.04$	0.95	0.61	1.56
$\phi = 0.97$	0.47	3.54	0.13
$\sigma_\eta^2 = 0.04$	0.69	1.17	0.59
$\phi = 0.99$	0.18	74.36	0.00
$\sigma_\eta^2 = 0.04$	0.47	73.98	0.01

about the better performance of the QML estimator. Further, using the GMM estimator in these circumstances could lead to huge losses in efficiency. Observe, for example, the case when $\sigma_\eta^2 = 1$ and $\phi = 0.9$, in Table 7. The extremely big standard deviations in this case may be due to the fact that the excess kurtosis of y_t is big and, therefore, the sample moments are not reliable estimators of the population moments.

5. An empirical example

In this section, we illustrate the QML estimation method by fitting SV models to the yen/dollar exchange rate. The data consist of daily observations of weekdays close exchange rates from 1/10/81 to 28/6/85, giving $T = 946$ (this data was also used in the empirical application reported in Harvey et al., 1992). The analysed series is the first differences of the logarithms of the spot price, i.e., the rates of return. For convenience, the rates of return have been centered about the sample mean prior to analysis.

Table 8

Empirical estimates of SV models using daily yen/dollar exchange rates with asymptotic standard deviations in parentheses

(a) Stationary SV model			
	$\varepsilon_t \sim N(0,1)$		$\varepsilon_t \sim$ Student-t
	QML	GMM	QML
γ	− 0.0551	− 0.3045	− 0.0493
ϕ	0.9948	0.9727	0.9954
	(0.0046)	(0.0942)	(0.0043)
σ_η^2	0.0048	0.1627	0.0042
	(0.0034)	(0.0229)	(0.0031)
σ_ξ^2	−	−	5.2989
			(0.4046)
ln L	− 1272.64	−	− 1271.5015

(b) Random walk volatility model			
	$\varepsilon_t \sim N(0,1)$		$\varepsilon_t \sim$ Student-t
	QML	MM	QML
σ_η^2	0.0034	0.7649	0.0030
	(0.0019)	(0.6671)	(0.0018)
σ_ξ^2	−	−	5.3039
			(0.4038)
log L	− 1273.64	−	− 1272.44

The Box–Ljung statistic for the data, y_t, based on ten lags, is 16.92, and therefore not significant at the 5% significance level. However, the Box–Ljung statistic for y_t^2 is 109.79, a highly significant value. Therefore, the dynamic properties of the yen/dollar exchange rate show up in the squares and not in the level of the series.

First, the stationary SV model in (2) is fitted to y_t. Table 8 shows the QML and GMM estimates when ε_t is assumed to be a standard normal variate. Both estimates of the autoregressive parameter imply persistence of the volatility and, as we have previously seen, in this circumstances the efficiency of the GMM method can be very low relative to the efficiency of the QML method. Therefore, we only consider the QML estimates.

As the QML estimate of the autoregressive parameter is very close to one, we estimate the random walk specification in (4). Once more we estimate by QML and using the method of moments (MM) estimator described in Section 4.1. The efficiency of the MM estimator is extremely low relative to the QML estimator. Moreover, the MM estimate is extremely big and not very reliable. Therefore, we concentrate on the QML estimates.

The random walk specification fits almost as well as the stationary SV specification. The Box–Ljung statistic of the innovations for ten lags has a value of 8.45, giving no indication of residual serial correlation. Fig. 1 shows the
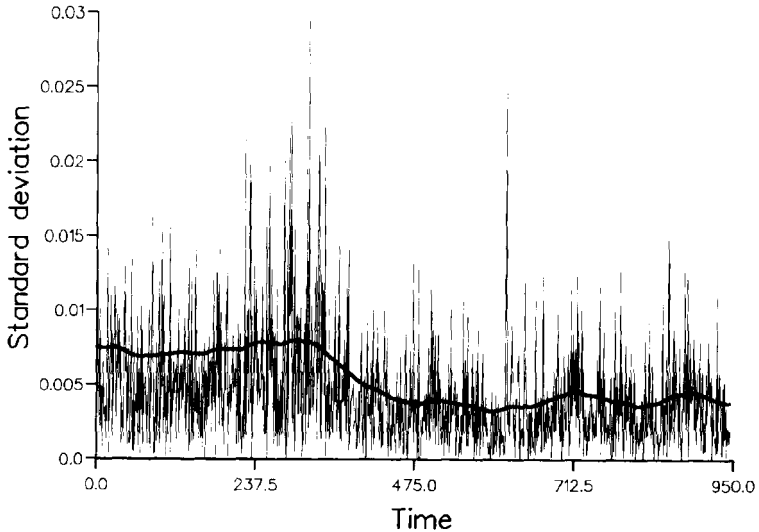

Fig. 1. Absolute value of first differences of logged yen/dollar exchange rate and smoothed estimates of volatility.

13

absolute values, $|y_t|$, together with the estimated volatility, $\exp\{0.5\,\tilde{h}_{t/T}\}$, where $\tilde{h}_{t/T}$ is the MMSLE of the log-volatility level, h_t, as given by a smoothing algorithm.

Next, the normality assumption for ε_t is relaxed, and the variance of $\ln(\varepsilon_t^2)$, σ_ξ^2, is estimated as another parameter which is restricted to be greater than or equal to $\pi^2/2$. The estimation results also appear in Table 8.

The number of degrees of freedom implied by $\hat{\sigma}_\xi^2$ in the random walk case is 6.35. The LR test statistic takes a value of 2.4. The critical value for $\alpha = 10\%$ is 1.642. Therefore, we reject normality at the 10% significance level. The evidence against normality is very weak but, as we have previously seen, the efficiency of $\hat{\sigma}_\eta^2$ is greater if we also estimate σ_ξ^2. Consequently, we maintain as our final estimate $\hat{\sigma}_\eta^2 = 0.0030$. Monte Carlo experiments carried out for these parameter values have shown that in this case, the ASE is a good approximation to the finite-sample standard error.

6. Conclusions

The QML estimator of the parameters of SV models is easy to implement and has good finite-sample properties. It is shown that for the parameter values often found in empirical analysis of high-frequency financial time series, the QML estimator outperforms in terms of efficiency some estimators based on the GMM principle.

The extension to heavy-tailed distributions can be carried out very easily using the t-distribution. In this case, the estimation of SV models can be carried out by assuming the number of degrees of freedom, v, is known or estimating it as another parameter. When comparing the asymptotic variances of the QML estimators obtained by assuming that v is known or when it is estimated, we found the somewhat counterintuitive result that there is a slight loss in efficiency when estimating σ_η^2 in the former case. This result may be due to the fact that since our estimators are QML, they are not fully efficient. In any case, given that assuming an incorrect value of v leads to an inconsistent estimate of σ_η^2, there seems to be no reason for making assumptions about the distribution of ε_t.

Appendix

Asymptotic variance and covariance matrices of the QML estimator of some stochastic volatility models

The derivation of the analytic expressions of all the asymptotic variance and covariance matrices in this article has been carried out, using the results in Dunsmuir (1979), in an appendix which is available upon request.

The asymptotic variance and covariance matrix of the QML estimator of SV models is given by $C(\Psi) = 2A^{-1} + A^{-1}BA^{-1}$, with

$$A = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\partial \log g(\lambda)}{\partial \Psi} \frac{\partial \log g(\lambda)}{\partial \Psi'} \, d\lambda,$$

where Ψ is the vector of parameters to be estimated, $g(\lambda)$ is the spectral generating function of the stationary form of $\log y_t^2$, and

$$B_{ij} = \kappa \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} k^*(\lambda) \frac{\partial g^{-1}(\lambda)}{\partial \Psi_i} k(\lambda)\, d\lambda \right]_{22} \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} k^*(\lambda) \frac{\partial g^{-1}(\lambda)}{\partial \Psi_j} k(\lambda)\, d\lambda \right]_{22},$$

where κ is the fourth cumulant of $\log \varepsilon_t^2$, $k(\lambda) = \sum_0^\infty D_j e^{i\lambda j}$, D_j is the 1×2 vector of coefficients corresponding to $(\eta_t\ \xi_t)'$ in the stationary representation of $\log y_t^2$, $k^*(\lambda)$ is the transpose conjugate of $k(\lambda)$, and $[M]_{uv}$ denotes the $\{uv\}$th element of matrix M.

(a) *Stationary SV model with Gaussian errors*

A consists of the following elements:

$$a_{11} = 4\left[\frac{1}{2(1-\phi^2)} + \frac{\phi^2 a}{(a^2-b^2)^{3/2}} + \frac{1}{b^2}\left(1 + \frac{a^3}{(a^2-b^2)^{3/2}} - \frac{2a}{(a^2-b^2)^{1/2}}\right)\right.$$

$$+ \frac{2\phi b}{(a^2-b^2)^{3/2}} - \frac{1}{q} + \frac{2}{q}\left(\frac{\phi^2}{(a^2-b^2)^{1/2}} + \frac{a}{b^2}\left(\frac{a}{(a^2-b^2)^{1/2}} - 1\right)\right.$$

$$\left.\left.+ \left(1 - \frac{a}{(a^2-b^2)^{1/2}}\right)\right)\right].$$

$$a_{12} = \frac{4}{\pi^2}\left[\frac{1}{q}\left(\frac{\phi}{(a^2-b^2)^{1/2}} - \frac{1}{b}\left(1 - \frac{a}{(a^2-b^2)^{1/2}}\right)\right)\right.$$

$$\left. + \frac{\phi a}{(a^2-b^2)^{3/2}} + \frac{b}{(a^2-b^2)^{3/2}}\right],$$

$$a_{22} = \frac{4a}{\pi^4(a^2-b^2)^{3/2}},$$

where $a = q + 1 + \phi^2$, $b = -2\phi$, and $q = 2\sigma_\eta^2/\pi^2$.

B is given by

$$B = \frac{16}{(a^2 - b^2)^3}$$

$$\times \begin{bmatrix} q^2(\phi a + b)^2 & \dfrac{-q((1 + \phi^2)a + 2\phi b)(\phi a + b)}{\pi^2} \\[2ex] \dfrac{-q((1 + \phi^2)a + 2\phi b)(\phi a + b)}{\pi^2} & \dfrac{((1 + \phi^2)a + 2\phi b)^2}{\pi^4} \end{bmatrix}.$$

(b) *Random walk volatility model with Gaussian errors*

$$C(\sigma_\eta^2) = \frac{2}{\sigma_\eta^2}\left[(\sigma_\eta^4 + 2\sigma_\eta^2\pi^2)^{3/2} + \frac{2\pi^4\sigma_\eta^4}{\sigma_\eta^2 + 2\pi^2}\right].$$

(c) *Stationary SV model with Student-t errors*

A consists of the following elements:

$$a_{11} = \frac{1}{\sigma^4}\left[\frac{(1 + \phi^2)^2 a}{(a^2 - b^2)^{3/2}} + 1 + \frac{a^3}{(a^2 - b^2)^{3/2}} - \frac{2a}{(a^2 - b^2)^{1/2}} - \frac{8\phi^2(1 + \phi^2)}{(a^2 - b^2)^{3/2}}\right],$$

$$a_{12} = \frac{-2q}{\sigma^2}\left[\frac{\phi a + b}{(a^2 - b^2)^{3/2}}\right].$$

$$a_{13} = \frac{1}{\sigma^4}\left[\frac{1 + q + \phi^2(q - 2 + \phi^2)}{(a^2 - b^2)^{3/2}}\right],$$

$$a_{22} = 4\left[\frac{1}{2(1 - \phi^2)} + \frac{\phi^2 a}{(a^2 - b^2)^{3/2}} + \frac{1}{b^2}\left(1 + \frac{a^3}{(a^2 - b^2)^{3/2}} - \frac{2a}{(a^2 - b^2)^{1/2}}\right)\right.$$

$$+ \frac{2\phi b}{(a^2 - b^2)^{3/2}} - \frac{1}{q} + \frac{2}{q}\left(\frac{\phi^2}{(a^2 - b^2)^{1/2}} + \frac{a}{b^2}\left(\frac{a}{(a^2 - b^2)^{1/2}} - 1\right)\right)$$

$$+ \left.\left(1 - \frac{a}{(a^2 - b^2)^{1/2}}\right)\right)\bigg],$$

$$a_{23} = \frac{2}{\sigma^2}\left[\frac{1}{q}\left(\frac{\phi}{(a^2 - b^2)^{1/2}} - \frac{1}{b}\left(1 - \frac{a}{(a^2 - b^2)^{1/2}}\right)\right)\right.$$

$$+ \left.\frac{\phi a}{(a^2 - b^2)^{3/2}} + \frac{b}{(a^2 - b^2)^{3/2}}\right],$$

$$a_{33} = \frac{a}{\sigma^4(a^2 - b^2)^{3/2}}.$$

16

B is given by

$$B = \frac{\pi^4 + \psi^{(3)}(v/2)}{\sigma^4(a^2 - b^2)^3} b_1 b_1',$$

where $\sigma^2 = \pi^2/2 + \psi'(v/2)$, $\psi^{(3)}(\cdot)$ is the Tetragamma function and

$$b_1 = \begin{bmatrix} \dfrac{(1 + \phi^2)a + (a^2 - b^2)^{3/2} - a^3 + 2ab^2 + 4\phi b(1 + \phi^2)}{\sigma^2} \\ - 2q(\phi a + b) \\ \dfrac{(1 + \phi^2)a + 2\phi b}{\sigma^2} \end{bmatrix}.$$

(d) *Random walk volatility model with Student-t errors*

A^{-1} is given by

$$A^{-1} = \frac{\sigma^4}{(q + 2) - (q(q + 4))^{1/2}} \begin{bmatrix} (q + 2) & - 2q \\ - 2q & (q(q + 4))^{3/2} - q^2(q + 6) \end{bmatrix}.$$

B is given by

$$B = \frac{4(\pi^4 + \psi^{(3)}(v/2)}{\sigma^8(q(q + 4))^3}$$
$$\times \begin{bmatrix} 0.25((q^2 + 4q)^{3/2} - q^2(q + 6))^2 & 0.59((q^2 + 4q)^{3/2} - q^2(q + 6)) \\ 0.59((q^2 + 4q)^{3/2} - q^2(q + 6)) & q^2 \end{bmatrix}.$$

References

Abramovitz, M. and N. Stegun, 1970, Handbook of mathematical functions (Dover Publications, New York, NY).

Andersen, T.G., 1992, Volatility, Unpublished manuscript (Northwestern University, Evanston, IL).

Bollerslev, T., 1987, A conditional heteroskedastic time series model for speculative prices and rates of return, Review of Economics and Statistics 69, 542–547.

Bollerslev, T., R.Y. Chou, and K.F. Kroner, 1992, ARCH modeling in finance: A review of the theory and empirical evidence, Journal of Econometrics 52, 5–59.

Chesney, M. and L. Scott, 1989, Pricing European currency options: A comparison of the modified Black–Scholes model and a random variance model, Journal of Financial and Quantitative Analysis 24, 267–284.

Dunsmuir, W., 1979, A central limit theorem for parameter estimation in stationary time series and its applications to models for a signal observed white noise, Annals of Statistics 7, 490–506.

Engle, R.F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of the U.K. inflation, Econometrica 50, 987–1007.

Hansen, L.P., 1982, Large sample properties of generalized method of moments estimators, Econometrica 50, 1029–1054.

Harvey, A.C., 1989, Forecasting, structural models and the Kalman filter (Cambridge University Press, Cambridge).

Harvey, A.C., E. Ruiz, and N.G. Shephard, 1992, Multivariate stochastic variance model, Financial Markets Group discussion paper (London School of Economics, London).

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

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

Taylor, S.J., 1991, Modeling stochastic volatility, Unpublished manuscript (Lancaster University, Lancaster).

Vetzal, K., 1991, Stochastic short rate volatility and the pricing of bonds and bond options, Unpublished Ph.D. dissertation (University of Toronto, Toronto).