Asymptotic Results for GMM Estimators of Stochastic Volatility Models

Geert Dhaene*
K.U.Leuven

Olivier Vergote
K.U.Leuven

May 2003

Abstract

We derive closed-form expressions for the optimal weighting matrix for GMM estimation of the stochastic volatility model with AR(1) log-volatility, and for the asymptotic covariance matrix of the resulting estimator. The moment conditions considered are generated by the absolute observations (which is the standard approach in this literature) or by the log-squared observations. We use the expressions to compare the performances of GMM and other estimators that have been proposed, and to optimally select small sets of moment conditions from very large sets.

JEL: C13, C22
Keywords: Stochastic volatility, GMM

*Corresponding author. Address: K.U.Leuven, Department of Economics, Naamsestraat 69, B-3000 Leuven, Belgium. Tel. +32 16 326798; Fax +32 16 326796; Email: geert.dhaene@econ.kuleuven.ac.be.

1 Introduction

Over the last two decades there has been an increasing interest in stochastic volatility (SV), which was introduced by Clark (1973) and extended by Tauchen and Pitts (1983), as a framework for the analysis of time-varying volatility in financial markets. This interest is partly due to an important contribution by Hull and White (1987), where SV models arise as discrete time approximations to continuous time volatility diffusions used in option pricing. More generally, it is recognized that SV models constitute a valuable alternative to GARCH-type models for analysing financial time series (Ghysels, Harvey, and Renault (1996), Shephard (1996)).

Due to the fact that in SV models the mean and the volatility are driven by separate stochastic processes (implying that volatility is unobservable), SV models are much harder to estimate than GARCH models. This paper presents analytical results that may be used to improve and assess the quality of GMM-based estimation of SV models. GMM, while not asymptotically efficient, is still the simplest estimation method for SV models currently available. It has been proposed by Taylor (1986) and Melino and Turnbull (1990), and its properties have been studied using Monte Carlo methods by Jacquier, Polson, and Rossi (1994), Andersen and Sørensen (1996, 1997), and Andersen, Chung, and Sørensen (1999). Other available estimation methods for SV models include quasi-maximum likelihood (Nelson (1988), Harvey, Ruiz, and Shephard (1994), Ruiz (1994)), simulated maximum likelihood (Danielsson and Richard (1993), Danielsson (1994)), simulation-based GMM (Duffie and Singleton (1993)), indirect inference (Gouriéroux, Monfort, and Renault (1993), Monfardini (1998)), Markov chain Monte Carlo methods (Jacquier, Polson, and Rossi (1994), Kim, Shephard, and Chib (1998), Chib, Nardari, and Shephard (2002)), efficient method of moments (Gallant, Hsieh, and Tauchen (1997), Andersen, Chung, and

Sørensen (1999)), Monte Carlo maximum likelihood (Sandmann and Koopman (1998)), and (approximate) maximum likelihood (Fridman and Harris (1998)). Apart from quasi-maximum likelihood, all of these methods are computationally more demanding, as they rely – often quite heavily – on numerical simulation and/or integration techniques both for obtaining point estimates and for assessing the accuracy of the latter. In view of its simplicity, we consider GMM estimation as a useful alternative to the more elaborate methods.

In this paper we derive closed-form expressions for the optimal weighting matrix for GMM estimation of the basic SV model, and for the asymptotic covariance matrix of the optimal GMM estimator, for a large class of moment conditions. To date, applications of GMM in this context have typically relied on a nonparametrically estimated weighting matrix, because an expression for the optimal weighting matrix (as a function of the parameters) was not available.

The moment conditions that we consider fall into two categories. The first set of conditions is obtained by considering the first two moments and the autocovariances of any order of the log-squared observations. These conditions have recently been considered by Wright (1999), in connection with the fractionally integrated SV model. The second set of moment conditions are derived from the absolute observations and are more standard in this literature. We study moment conditions that involve the product of any number of absolute observations, each one raised to any positive integer power and lagged any number of periods. This set considerably extends the set of moment conditions that have been employed so far. The results that we present pertain to any selection of moment conditions from these two sets.

In Section 2, we present the basic SV model and the moment conditions. Expressions for the optimal weighting matrix and the asymptotic covariance matrix of the GMM estimator are derived in Section 3. Section 4 presents some comparative evidence on the relative efficiencies of the GMM and other estimators (partly compiled from the literature). We also show how the analytical results of this paper provide a fast and accurate tool to select a small set of highly informative moment conditions from very large sets of moment conditions. Section 5 concludes. Proofs are given in the Appendix.

2 Moment conditions for the SV model

The basic SV model is given by

$$y_t = \exp\left(h_t/2\right) u_t, \tag{1}$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \sigma\sqrt{1-\phi^2}v_t, \tag{2}$$

where y_t is observable, h_t is latent log-volatility, (u_t, v_t) is i.i.d. $N(0, I)$, and $\theta = (\mu, \phi, \sigma)'$ is a vector of parameters. The restriction $|\phi| < 1$ is imposed, ensuring that y_t is stationary and ergodic. While it is more common to parameterise the model in terms of $\lambda = (\alpha, \phi, \omega)'$, with $\alpha = \mu(1 - \phi)$ and $\omega = \sigma\sqrt{1-\phi^2}$, we prefer the parameterisation in terms of θ for algebraic reasons and because of an invariance with respect to μ given below. For comparison with earlier studies, however, numerical standard errors will be presented in terms of λ.

From the point of view of inference, the fundamental problem with the SV model is the latent character of h_t, which makes it difficult to compute the values of the likelihood function and hence to estimate θ by maximum likelihood. It is easy, however, to derive moment conditions implied by the SV model and then to apply the Generalized Method of Moments (Hansen, 1982). The moment conditions considered in this paper relate either to the log-squared observations, $\log y_t^2$, or to the absolute observations, $|y_t|$. The latter class of moment conditions constitutes the standard approach to GMM estimation of SV models (Taylor (1986), Melino and Turnbull (1990), Jacquier, Polson, and Rossi (1994), Andersen and Sørensen (1996, 1997), Andersen, Chung, and Sørensen (1999)). The former class of moment conditions is suggested in passing by Jacquier, Polson, and Rossi (1994), and is effectively employed by Wright (1999) in the context of the fractionally integrated SV model.

Moment conditions related to $\log y_t^2$ are easily obtained. It follows from (1) that $\log y_t^2 = h_t + \log u_t^2$. The mean and variance of $\log u_t^2$ are known to be $c_1 = -\log 2 - \gamma = -1.2704$ and $c_2 = \frac{1}{2}\pi^2 = 4.9348$, respectively, where $\gamma = 0.5772$ is Euler's constant. Let

$$\begin{aligned} z_t &= \log y_t^2 - \mu - c_1 \\ &= h_t - \mu + \log u_t^2 - c_1. \end{aligned}$$

Since $h_t \sim N(\mu, \sigma^2)$, $\mathrm{Cov}(h_t, h_{t-i}) = \phi^{|i|}\sigma^2$, and u_t is i.i.d. and independent of h_t, it follows that

$$E\left[z_t\right] = 0, \tag{3}$$

$$E\left[z_t z_{t-i}\right] = \phi^i \sigma^2 + I_{(i=0)} c_2, \qquad i \geq 0, \tag{4}$$

where $I_{(\cdot)}$ is the indicator function. It can be shown that none of these conditions is redundant in the sense of Breusch et al. (1999).

We now derive the class of moment conditions generated by the expectation of $|y_{t_1}^{i_1} ... y_{t_p}^{i_p}|$, where $i_1, ..., i_p$ are positive integers and $t_1 > ... > t_p$. Let ν_i be the i-th absolute moment of a standard normal random variate, i.e.

$$\nu_i = E|u_t|^i = \frac{2^{i/2}}{\sqrt{\pi}} \Gamma\left(\frac{i+1}{2}\right) = \begin{cases} \sqrt{\frac{2^i}{\pi}}\left(\frac{i-1}{2}\right)! & \text{for } i \text{ odd}, \\ \frac{i!}{(i/2)! 2^{i/2}} & \text{for } i \text{ even}, \end{cases}$$

where $\Gamma(z)$ is the gamma function. Then, because $t_1 > ... > t_p$,

$$E \prod_{j=1}^{p} \left| \nu_{i_j}^{-1} u_{t_j}^{i_j} \right| = 1.$$

Furthermore, $\sum_{j=1}^{p} i_j h_{t_j}$ is normally distributed with mean $\mu \sum_{j=1}^{p} i_j$ and variance $\sigma^2 \sum_{j,j'=1}^{p} i_j i_{j'} \phi^{|t_j - t_{j'}|}$. So, by property that $E\exp(X) = \exp(a + \frac{1}{2}b^2)$

when $X \sim N(a, b^2)$, we have

$$E \exp\left(\frac{1}{2}\sum_{j=1}^{p} i_j h_{t_j}\right) = \exp\left(\delta_{t_1,...,t_p}^{i_1,...,i_p}\right),$$

where

$$\delta_{t_1,...,t_p}^{i_1,...,i_p} = \frac{\mu}{2}\sum_{j=1}^{p} i_j + \frac{\sigma^2}{8}\sum_{j,j'=1}^{p} i_j i_{j'} \phi^{|t_j - t_{j'}|}.$$

Hence, defining

$$\begin{aligned} Y_{t_1,...,t_p}^{i_1,...,i_p} &= \exp\left(-\delta_{t_1,...,t_p}^{i_1,...,i_p}\right)\prod_{j=1}^{p}\left|\nu_{i_j}^{-1} y_{t_j}^{i_j}\right| \\ &= \exp\left(-\delta_{t_1,...,t_p}^{i_1,...,i_p}\right)\exp\left(\frac{1}{2}\sum_{j=1}^{p} i_j h_{t_j}\right)\prod_{j=1}^{p}\left|\nu_{i_j}^{-1} u_{t_j}^{i_j}\right|, \end{aligned}$$

it follows that

$$E\left[Y_{t_1,...,t_p}^{i_1,...,i_p}\right] = 1, \qquad i_1, ..., i_p \geq 1; \quad t_1 > ... > t_p. \tag{5}$$

It is obvious that adding the same integer to $t_1, ..., t_p$ yields the same moment condition. As far as we know, within the class of moment conditions defined by (5), only moment conditions where $p = 1$ or where $p = 2$ and $i_1 = i_2 \in \{1, 2\}$ have so far been considered in the literature.

3 Optimal GMM

Let $E(f_t) = f$ be a finite selection of the set of moment conditions given by (3)–(5) that identifies θ. Let $g_t = f_t - f$. By assumption, the observations on y_t permit us to calculate $g_1, ..., g_T$ as functions of θ. The optimal GMM estimator (Hansen (1982)) of θ based on this selection is $\hat{\theta} = \arg\min_\theta \bar{g}'\hat{V}^{-1}\bar{g}$, where $\bar{g} = T^{-1}\sum_{t=1}^{T} g_t$ and $\hat{V}$ consistently estimates V, where

$$V = \sum_{l=-\infty}^{\infty} E(g_t g_{t-l}') = \sum_{l=-\infty}^{\infty} \text{Cov}(f_t, f_{t-l}').$$

The asymptotic covariance matrix of $\sqrt{T}(\hat{\theta}-\theta)$ is $(D'V^{-1}D)^{-1}$, where $D = E(\frac{\partial g_t}{\partial \theta'})$. Expressions for D and V, for an arbitrary selection of moment conditions, are presented below. These expressions make it possible to compute the optimal weighting matrix V^{-1} and the asymptotic covariance matrix $(D'V^{-1}D)^{-1}$ of the GMM estimator as functions of the parameter values. Substituting estimates for these parameter values yields estimates of V^{-1} and $(D'V^{-1}D)^{-1}$, which will generally be more precise than the nonparametric estimator based on Bartlett weights that is routinely used in a GMM context. The Monte Carlo results of Andersen and Sørensen (1996) show that the latter estimator of V may be imprecise even in samples of size $50,000$. Using the expressions presented here avoids such problems. Furthermore, the expression for V makes it also possible to estimate θ by the continuous-updating GMM estimator of Hansen, Heaton, and Yaron (1996), that is, by solving $\min_\theta \bar{g}'V^{-1}\bar{g}$.

Some straightforward calculus shows that the rows of D are to be selected (according to the selection of moments) from

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -i\phi^{i-1}\sigma^2 & -2\phi^i\sigma \\ -\nabla_\mu \delta^{i_1,\ldots,i_p}_{t_1,\ldots,t_p} & -\nabla_\phi \delta^{i_1,\ldots,i_p}_{t_1,\ldots,t_p} & -\nabla_\sigma \delta^{i_1,\ldots,i_p}_{t_1,\ldots,t_p} \end{pmatrix} \begin{matrix} \\ i \geq 0; \\ i_1,\ldots,i_p \geq 1; \quad t_1 > \ldots > t_p; \end{matrix}$$

where

$$\nabla_\mu \delta^{i_1,\ldots,i_p}_{t_1,\ldots,t_p} = \frac{1}{2}\sum_{j=1}^{p} i_j,$$

$$\nabla_\phi \delta^{i_1,\ldots,i_p}_{t_1,\ldots,t_p} = \frac{\sigma^2}{8}\sum_{j,j'=1}^{p} i_j i_{j'} |t_j - t_{j'}| \phi^{|t_j - t_{j'}|-1},$$

$$\nabla_\sigma \delta^{i_1,\ldots,i_p}_{t_1,\ldots,t_p} = \frac{\sigma}{4}\sum_{j,j'=1}^{p} i_j i_{j'} \phi^{|t_j - t_{j'}|}.$$

The main result of this paper is an expression for the elements of V, given in Theorem 1 below. Let $c_i = E(\log u_t^2 - c_1)^i$, $i = 3, 4$. It is shown in the Appendix that

$$c_3 = -14\zeta(3) = -16.829,$$
$$c_4 = \tfrac{7}{4}\pi^4 = 170.47,$$

where $\zeta(z)$ is the Riemann zeta function. For $i \geq 1$, let

$$\kappa_i = \log 2 + \psi\left(\frac{i+1}{2}\right) - c_1,$$
$$\xi_i = \kappa_i^2 + \psi'\left(\frac{i+1}{2}\right) - c_2,$$

where $\psi(z) = \frac{d}{dz}\log\Gamma(z)$, the digamma function, and $\psi'(z) = \frac{d}{dz}\psi(z)$, the trigamma function.

Theorem 1 *For any a_t and b_t, let $V(a_t, b_t) = \sum_{l=-\infty}^{\infty} \mathrm{Cov}(a_t, b_{t-l})$. Let $i, j \geq 0$. Then*

$$V(z_t, z_t) = \frac{1+\phi}{1-\phi}\sigma^2 + c_2, \tag{6}$$
$$V(z_t, z_t z_{t-j}) = I_{(j=0)} c_3, \tag{7}$$
$$V(z_t z_{t-i}, z_t z_{t-j}) = A_1\sigma^4 + A_2 c_2 \sigma^2 + I_{(i=j\neq 0)} c_2^2 + I_{(i=j=0)}(c_4 - c_2^2), \tag{8}$$

where

$$A_1 = |i-j|\,\phi^{|i-j|} + |i+j|\,\phi^{|i+j|} + \left(\phi^{|i-j|} + \phi^{|i+j|}\right)\frac{1+\phi^2}{1-\phi^2},$$
$$A_2 = 2\left(\phi^{|i-j|} + \phi^{|i+j|}\right).$$

Let $i_1, ..., i_{p+q} \geq 1$; $t_1 > ... > t_p$; $t_{p+1} > ... > t_{p+q}$; and let

$$L = \{l | \{t_1, ..., t_p\} \cap \{t_{p+1} - l, ..., t_{p+q} - l\} \neq \varnothing\}.$$

Then

$$V\left(Y_{t_1,...,t_p}^{i_1,...,i_p}, Y_{t_{p+1},...,t_{p+q}}^{i_{p+1},...,i_{p+q}}\right) = B + \sum_{l\in L}(B_l + 1)C_l, \tag{9}$$

where

$$B = \sum_{l=-\infty}^{\infty} B_l, \qquad B_l = \exp\left(\frac{\sigma^2}{4}\sum_{j=1}^{p}\sum_{j'=p+1}^{p+q} i_j i_{j'} \phi^{|t_j - t_{j'} + l|}\right) - 1,$$

$$C_l = \left(\prod_{j=1}^{p+q} \nu_{i_j}^{-1}\right) E\left(\prod_{j=1}^{p}\left|u_{t_j}^{i_j}\right| \prod_{j=p+1}^{p+q}\left|u_{t_j-l}^{i_j}\right|\right) - 1,$$

and

$$V\left(z_t, Y_{t_1,\ldots,t_p}^{i_1,\ldots,i_p}\right) = D_1\sigma^2 + D_2, \tag{10}$$

$$V\left(z_t z_{t-i}, Y_{t_1,\ldots,t_p}^{i_1,\ldots,i_p}\right) = D_1'\sigma^4 + D_2'\sigma^2 + D_3', \tag{11}$$

where

$$D_1 = \frac{1}{2}\left(\frac{1+\phi}{1-\phi}\right)\sum_{j=1}^{p} i_j, \qquad D_2 = \sum_{j=1}^{p} \kappa_{i_j},$$

$$D_1' = \frac{1}{4}\sum_{j,j'=1}^{p} i_j i_{j'} \phi^{\left|t_j - t_{j'} + i\right|}\left(|t_j - t_{j'} + i| + \frac{1+\phi^2}{1-\phi^2}\right),$$

$$D_2' = \frac{1}{2}\sum_{j,j'=1}^{p} i_j \kappa_{i_{j'}}\left(\phi^{\left|t_j - t_{j'} + i\right|} + \phi^{\left|t_j - t_{j'} - i\right|}\right),$$

$$D_3' = I_{(i=0)}\sum_{j=1}^{p}\xi_{i_j} + \sum_{j,j'=1}^{p} I_{\left(i=t_j-t_{j'}\neq 0\right)}\kappa_{i_j}\kappa_{i_{j'}}.$$

We see that, not unexpectedly, the optimal weighting matrix, V^{-1}, and the GMM asymptotic covariance matrix, $(D'V^{-1}D)^{-1}$, do not depend on μ. From a computational point of view, notice that L has at most pq elements, so computing $\sum_{l\in L}(B_l+1)C_l$ requires a finite number of steps. Furthermore, B can be approximated by $B(I) = \sum_{l=-I}^{I} B_l$, where I is a positive integer. As the following lemma shows, the error of approximation $|B - B(I)|$ is bounded by an exponentially decaying function in I, and this bound can be inverted to determine I as a function of the desired accuracy of the approximation.

Lemma 1 *Let $i_1, ..., i_{p+q} \geq 1$; $t_1 > ... > t_p$; $t_{p+1} > ... > t_{p+q}$; and let I be a positive integer. Then*

$$|B - B(I)| \leq 2\frac{\exp(a|\phi|^I) - 1}{1 - |\phi|},$$

where

$$a = \frac{\sigma^2}{4}\sum_{j=1}^{p}\sum_{j'=p+1}^{p+q} i_j i_{j'} \, |\phi|^{-|t_j - t_{j'}|} \,.$$

If one is interested in λ rather than θ, one may apply the transformation $\theta \longmapsto \lambda(\theta)$ to yield $\hat{\lambda} = \lambda(\hat{\theta})$, the optimal GMM estimator of λ, which has asymptotic covariance matrix $(\frac{\partial\lambda}{\partial\theta'})(D'V^{-1}D)^{-1}(\frac{\partial\lambda}{\partial\theta'})'$, with

$$\frac{\partial\lambda}{\partial\theta'} = \begin{pmatrix} 1-\phi & -\mu & 0 \\ 0 & 1 & 0 \\ 0 & -\sigma\phi(1-\phi^2)^{-1/2} & (1-\phi^2)^{1/2} \end{pmatrix}.$$

4 Comparison of GMM and other estimators

In this section we first compare the relative efficiencies of GMM and other estimators, for two sets of values of λ, namely $(\alpha, \phi, \omega) = (-0.736, 0.90, 0.363)$ and $(\alpha, \phi, \omega) = (-0.1472, 0.98, 0.1657)$. These parameter values have been used in earlier Monte Carlo studies (Jacquier, Polson, and Rossi (1994), Andersen and Sørensen (1996), Fridman and Harris (1998), Sandmann and Koopman (1998), Andersen, Chung, and Sørensen (1999)). Tables 1 and 2 present the results. The asymptotic standard errors of the GMM estimators were computed using the expressions derived above. The moment conditions were selected from the set related to the log-squared observations, or from the set related to the absolute observations, or from both. For comparability with other studies, from (5) we only selected moment conditions for which $p = 1$ or for which $p = 2$ and $i_1 = i_2 \in \{1, 2\}$. The (finite sample) standard errors of the other estimators were taken from the aforementioned Monte Carlo studies and multiplied by $\sqrt{T}$. The relative asymptotic efficiency of the GMM estimators is seen to increase rapidly with the number of moments, at least when this number is small. Using a large number of moment conditions yields asymptotic standard errors slightly above those of the MCMC method, which is known to be asymptotically efficient. In this respect, it appears that some of the published standard errors regarding ML and Monte Carlo ML are not in line with those of the MCMC method.

In Monte Carlo studies it is often found that the small sample bias of the GMM estimator grows with the number of moment conditions. Newey and Smith (2000, 2001) show that the number of terms of the second-order bias increases linearly with the number of moment conditions. Thus, rather than using a large number of moment conditions (relative to the sample size), it is in terms of bias often safer to select only a small number of them. It is important, then, to choose the moments judiciously, in the sense that they contain as much information as possible for the estimand. Several authors have addressed the question of how to select the moment conditions to estimate the SV model, essentially by resorting to Monte Carlo simulation of the accuracy of the GMM estimator for any given choice of moments. The results of the previous section provide a more precise and much faster tool to guide the choice of moments. To illustrate this point, consider the sets M_L and M_A of log-moment and absolute moment conditions, respectively, defined as

$$M_L : \text{(3)–(4) with } i \le 50,$$

$$M_A : \text{(5) with } \max_{j,j'} |t_j - t_{j'}| \le 15 \text{ and } \textstyle\sum_{j=1}^{p} i_j \le \begin{cases} 20 & \text{for } p = 1, \\ 4 & \text{for } p = 2, 3, 4. \end{cases}$$

The sets M_L and M_A comprise 52 and 985 moment conditions, respectively. We performed a search for the set of k moment conditions, selected from either M_L, M_A, or $M_L \cup M_A$, that yield the smallest asymptotic standard error of $\hat{\phi}$. Global optimisation, by enumeration, was performed over M_L for $k = 3, 4, 5$, and over M_A and $M_L \cup M_A$ for $k = 3$. Global optimisation over M_A and $M_L \cup M_A$ for $k = 4$ and $k = 5$ turned out to be infeasible in terms of computation time, and

Table 1: Standard errors of $\sqrt{T}\hat{\lambda}$

T	# Moments	Method of estimation	Standard Error of		
			$\sqrt{T}\hat{\alpha}$	$\sqrt{T}\hat{\phi}$	$\sqrt{T}\hat{\omega}$
∞	3	GMM (log-moments)[a]	127.52	17.31	32.66
∞	12	GMM (log-moments)[a]	12.04	1.63	3.80
∞	27	GMM (log-moments)[a]	10.06	1.36	3.22
∞	102	GMM (log-moments)[a]	10.04	1.36	3.22
∞	3	GMM (absolute moments)[b]	178.46	24.18	46.78
∞	15	GMM (absolute moments)[b]	11.34	1.53	2.96
∞	30	GMM (absolute moments)[b]	8.14	1.10	2.18
∞	75	GMM (absolute moments)[b]	7.55	1.02	2.03
∞	14	GMM (joint moments)[c]	16.92	2.29	4.27
∞	22	GMM (joint moments)[c]	11.30	1.53	2.92
∞	42	GMM (joint moments)[c]	8.12	1.10	2.14
∞	102	GMM (joint moments)[c]	7.53	1.02	1.99
10000	14	Infeasible GMM[d] (true weight)	11.4	1.6	3.1
4000	14	Infeasible GMM[d] (true weight)	10.6	1.5	3.1
4000	4	EMM: GARCH(1,1)[e]	9.51	1.2	3.1
4000	6	EMM: GARCH(1,1) - Kz(2)[e]	9.68	1.3	3.2
4000	8	EMM: GARCH(1,1) - Kz(4)[e]	8.28	1.1	2.1
2000	24	GMM[f]	18	3	3.8
2000	-	Quasi-ML[f]	20	3	4.8
2000	-	MCMC[f]	6.6	1	1.5
500	-	ML[g]	9.1	1	2
500	-	Monte Carlo ML[h]	0.5	2.2	2

Parameter values: $(\alpha, \phi, \omega) = (-0.736, 0.90, 0.363)$.
GMM conditions are selected from Eqs. (3)–(5), as indicated below. Most footnotes refer to multiple lines in the Table.

a. Eqs. (3)–(4) with i running from 0 to $1, 10, 25$, and 100, respectively.

b. Eqs. (5) with $p = 1$, i_1 running from 1 to $1, 5, 10$, and 25 respectively; and Eq. (5) with $p = 2$, $i_1 = i_2 \in \{1, 2\}$, $t_1 - t_2$ running from 1 to $1, 5, 10$, and 25 respectively.

c. Eqs. (3)–(4) with i running from 0 to $3, 5, 10$, and 25, respectively; Eq. (5) with $p = 1$, i_1 running from 1 to $3, 5, 10$, and 25, respectively; and Eq. (5) with $p = 2$, $i_1 = i_2 \in \{1, 2\}$, $t_1 - t_2$ running from 1 to $3, 5, 10$, and 25, respectively.

d. Andersen and Sørensen (1996), Table 3: Eq. (5) with $p = 1$, i_1 running from 1 to 4; Eq. (5) with $p = 2$, $i_1 = i_2 = 1$, $t_1 - t_2 \in \{6, 8, 10, 12, 14\}$; and Eq. (5) with $p = 2$, $i_1 = i_2 = 2$, $t_1 - t_2 \in \{15, 17, 19, 21, 23\}$. 'Infeasible GMM' uses a nonparametric estimate of the weighting matrix based on a large sample of simulated data using true parameter values.

e. Andersen, Chung, and Sørensen (1999), Table 3.

f. Jacquier, Polson, and Rossi (1994), Tables 5–7. For GMM: Eq. (5) with $p = 1$, i_1 running from 1 to 4; and Eq. (5) with $p = 2$, $i_1 = i_2 \in \{1, 2\}$, $t_1 - t_2$ running from 1 to 10.

g. Fridman and Harris (1998), Table 1.

h. Sandmann and Koopman (1998), Table 2.

Table 2: Standard errors of $\sqrt{T}\hat{\lambda}$

T	# Moments	Method of estimation	Standard Error of		
			$\sqrt{T}\hat{\alpha}$	$\sqrt{T}\hat{\phi}$	$\sqrt{T}\hat{\omega}$
∞	3	GMM (log-moments)^a	136.37	18.53	77.30
∞	12	GMM (log-moments)^a	6.67	0.90	4.00
∞	27	GMM (log-moments)^a	2.96	0.40	1.71
∞	52	GMM (log-moments)^a	2.51	0.34	1.39
∞	102	GMM (log-moments)^a	2.49	0.34	1.37
∞	3	GMM (absolute moments)^b	264.71	35.95	150.79
∞	15	GMM (absolute moments)^b	8.49	1.15	4.79
∞	30	GMM (absolute moments)^b	4.15	0.56	2.28
∞	75	GMM (absolute moments)^b	2.48	0.34	1.23
∞	14	GMM (joint moments)^c	14.95	2.03	8.43
∞	22	GMM (joint moments)^c	8.45	1.15	4.76
∞	42	GMM (joint moments)^c	4.12	0.56	2.26
∞	102	GMM (joint moments)^c	2.44	0.33	1.20
4000	4	EMM: GARCH(1,1)^d	2.8	0.37	1.3
4000	4	EMM: GARCH(1,1) - Kz(2)^d	2.9	0.39	1.8
4000	6	EMM: GARCH(1,1) - Kz(4)^d	2.7	0.36	1.0
500	24	GMM^e	5.8	0.80	2
500	-	Quasi-ML^e	12	2	3.1
500	-	MCMC^e	2.7	0.4	1
500	-	ML^f	0.4	0.30	0.8
500	-	Monte Carlo ML^g	0.2	2	1

Parameter values: $(\alpha, \phi, \omega) = (-0.1472, 0.98, 0.1657)$.
GMM conditions are selected from Eqs. (3)–(5), as indicated below. Most footnotes refer to multiple lines in the Table.
a. Eqs. (3)–(4) with i running from 0 to $1, 10, 25, 50$, and 100, respectively.
b. Eq. (5) with $p = 1$, i_1 running from 1 to $1, 5, 10$, and 25 respectively; and Eq. (5) with $p = 2$, $i_1 = i_2 \in \{1, 2\}$, $t_1 - t_2$ running from 1 to $1, 5, 10$, and 25 respectively.
c. Eqs. (3)–(4) with i running from 0 to $3, 5, 10$, and 25, respectively; Eq. (5) with $p = 1$, i_1 running from 1 to $3, 5, 10$, and 25, respectively; and Eq. (5) with $p = 2$, $i_1 = i_2 \in \{1, 2\}$, $t_1 - t_2$ running from 1 to $3, 5, 10$, and 25, respectively.
d. Andersen, Chung, and Sørensen (1999), Table 3.
e. Jacquier, Polson, and Rossi (1994), Tables 5–7. For GMM: Eq. (5) with $p = 1$, i_1 running from 1 to 4; and Eq. (5) with $p = 2$, $i_1 = i_2 \in \{1, 2\}$, $t_1 - t_2$ running from 1 to 10.
f. Fridman and Harris (1998), Table 1.
g. Sandmann and Koopman (1998), Table 2.

Table 3: Asymptotic standard errors of $\sqrt{T}\hat{\phi}$ for parsimoniously selected moments

Moment set	k	Selected moments	Standard Error of $\sqrt{T}\hat{\alpha}$	$\sqrt{T}\hat{\phi}$	$\sqrt{T}\hat{\omega}$
M_L	3	$z_t; z_t z_{t-1}; z_t z_{t-11}$	18.31	2.49	5.41
M_A	3	$Y_t^2; Y_{t,t-7}^{1,2}; Y_{t,t-5,t-14}^{1,1,1}$	10.59	1.44	4.72
$M_L \cup M_A$	3	$z_t z_{t-10}; Y_t^2; Y_{t,t-7,t-15}^{1,1,1}$	10.08	1.37	4.07
M_L	4	$z_t; z_t z_{t-1}; z_t z_{t-10}; z_t z_{t-12}$	14.78	2.01	4.62
M_A	4	$Y_t^1; Y_t^2; Y_{t,t-10}^{1,1}; Y_{t,t-8,t-15}^{1,1,1}$	9.65	1.31	2.55
$M_L \cup M_A$	4	$z_t z_{t-10}; Y_t^2; Y_{t,t-5,t-14}^{1,1,1}; Y_{t,t-7,t-13}^{1,1,1}$	9.46	1.28	4.16
M_L	5	$z_t; z_t z_{t-1}; z_t z_{t-9}; z_t z_{t-11}; z_t z_{t-14}$	13.37	1.82	4.31
M_A	5	$Y_t^1; Y_t^2; Y_{t,t-7}^{1,1}; Y_{t,t-9,t-13}^{1,1}$	9.07	1.23	2.47
$M_L \cup M_A$	5	$Y_t^1; Y_t^2; Y_{t,t-7}^{1,1}; Y_{t,t-9,t-13}^{1,1}$	9.07	1.23	2.47

Parameter values: $(\alpha, \phi, \omega) = (-0.736, 0.90, 0.363)$.

in these cases we experimented with the Point Exchange algorithm (Fedorov (1972)). This algorithm does not necessarily yield the global optimum, and its output depends on the starting selection of moment conditions as input. By picking the starting selection at random and repeating this a couple of times, the algorithm was able to reproduce the global optimum in all cases where enumeration was possible. We therefore applied it in those cases where global optimisation was not feasible, without the guarantee of having found the globally optimal selection of moments from the specified sets. The parameter values were fixed at $(\alpha, \phi, \omega) = (-0.736, 0.90, 0.363)$, as in Table 1. Table 3 reports the selected moments and the asymptotic standard errors of the corresponding GMM estimators. Comparing Table 3 with Table 1 yields the following conclusions: (i) there is a dramatic increase in efficiency by selecting the moments in an optimal way; (ii) given that the efficiency bound for the asymptotic standard error of $\sqrt{T}\hat{\phi}$ (which is asymptotically attained by the MCMC estimator) appears to be around 1, the efficiency loss of the GMM estimator with optimal moment selection from $M_L \cup M_A$ is not excessively large, even in the just-identified case ($k = 3$); (iii) while M_A contains a richer (also a much larger) set of moment conditions than M_L – as is reflected by the smaller asymptotic standard errors – the combination of M_A and M_L may yield an improvement upon M_A, as is the case here for $k = 3$ and $k = 4$. Finally, we remark that the optimal selection of moment conditions from any given set generally depends on the parameter values.

5 Conclusion

The standard approach in the literature on GMM estimation of SV models has been to derive closed-form moment conditions from the expectations of $|y_t^i|$, $|y_{t_1} y_{t_2}|$ and $|y_{t_1}^2 y_{t_2}^2|$ for any i, t_1, and t_2. We have extended this class of conditions to include the expectation of $|y_{t_1}^{i_1} ... y_{t_p}^{i_p}|$ for arbitrary $i_1, ..., i_p$ and $t_1, ..., t_p$, and, following Wright (1999), the first two moments and the autocovariances of $\log y_t^2$. A closed-form expression for the optimal weighting matrix for any subset of those conditions has been derived and, as a by-product, an expression for the GMM asymptotic covariance matrix. These expressions can be used for improved GMM estimation of the SV model with AR(1) log-volatility and to compute GMM standard errors more accurately. It is also of interest to note that, upon redefining c_i, ν_i, κ_i, and ξ_i appropriately, all expressions are generalised to SV models where the multiplicative shocks in the mean equation (1) are non-normal.

Appendix

Calculation of c_3 and c_4. For any positive integer n, upon substituting $t = x^2/2$,

$$\begin{aligned}
\int_{-\infty}^{\infty} \left(\log \tfrac{x^2}{2}\right)^n \left(\tfrac{1}{\sqrt{2\pi}}\right) e^{-x^2/2} dx &= 2\int_{0}^{\infty} \left(\log \tfrac{x^2}{2}\right)^n \left(\tfrac{1}{\sqrt{2\pi}}\right) e^{-x^2/2} dx \\
&= \tfrac{1}{\sqrt{\pi}} \int_{0}^{\infty} (\log t)^n \, t^{-1/2} e^{-t} dt \\
&= \frac{\Gamma^{(n)}\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1}{2}\right)},
\end{aligned}$$

where $\Gamma^{(n)}(z)$ is the n-th derivative of $\Gamma(z)$. See Abramowitz and Stegun (1970) for properties and values of the gamma and related functions that are used below. Now, $c_1 = -\log 2 - \gamma = \psi\left(\frac{1}{2}\right) + \log 2$, where $\psi(z) = \frac{d}{dz}\log\Gamma(z) = \frac{\Gamma^{(1)}(z)}{\Gamma(z)}$. Hence, for $n = 3, 4$,

$$\begin{aligned}
c_n &= \int_{-\infty}^{\infty} (\log x^2 - c_1)^n \left(\tfrac{1}{\sqrt{2\pi}}\right) e^{-x^2/2} dx \\
&= \int_{-\infty}^{\infty} \left(\log \tfrac{x^2}{2} - \psi\left(\tfrac{1}{2}\right)\right)^n \left(\tfrac{1}{\sqrt{2\pi}}\right) e^{-x^2/2} dx \\
&= g_n\left(\tfrac{1}{2}\right),
\end{aligned}$$

where

$$g_n(z) = \sum_{i=0}^{n} \binom{n}{i} \frac{\Gamma^{(i)}(z)}{\Gamma(z)} \left(-\psi(z)\right)^{n-i}. \tag{12}$$

Taking successive derivatives of $\Gamma^{(1)}(z) = \Gamma(z)\,\psi(z)$ gives, upon rewriting,

$$\begin{aligned}
\Gamma^{(2)}(z) &= \Gamma(z)\left\{\psi'(z) + [\psi(z)]^2\right\}, \\
\Gamma^{(3)}(z) &= \Gamma(z)\left\{\psi''(z) + 3\psi'(z)\,\psi(z) + [\psi(z)]^3\right\}, \\
\Gamma^{(4)}(z) &= \Gamma(z)\left\{\psi'''(z) + 4\psi''(z) + 3\psi'(z)\left[\psi'(z) + 2[\psi(z)]^2\right] + [\psi(z)]^3\right\},
\end{aligned}$$

where primes denote derivatives. Substituting these expressions into (12) yields

$$g_3(z) = \psi''(z)$$

and

$$g_4(z) = \psi'''(z) + 3[\psi'(z)]^2.$$

Now, $\psi'\left(\frac{1}{2}\right)=\frac{\pi^2}{2}$, $\psi''\left(\frac{1}{2}\right)=-14\zeta(3)$, and $\psi'''\left(\frac{1}{2}\right)=\pi^4$, where $\zeta(3)=1.202$. Hence $c_3=-14\zeta\left(3\right)=-16.83$ and $c_4=\frac{7}{4}\pi^4=170.5$.

The proof of Theorem 1 makes use of the following lemmas.

Lemma 2 *Let $X\sim N(0,1)$ and let a be a positive integer. Then*

$$\mathrm{Cov}(\log X^2,|X^a|)=\nu_a\kappa_a \tag{13}$$

and

$$\mathrm{Cov}((\log X^2-c_1)^2,|X^a|)=\nu_a\xi_a. \tag{14}$$

Proof. Upon substituting $z=x^2/2$,

$$\begin{aligned}\mathrm{Cov}(\log X^2,|X^a|)&=2\int_0^\infty\left(\log x^2-c_1\right)x^a\left(\tfrac{1}{\sqrt{2\pi}}\right)e^{-x^2/2}dx\\&=\frac{2^{a/2}}{\sqrt{\pi}}\int_0^\infty\left(\log z+\log 2-c_1\right)z^{(a-1)/2}e^{-z}dz\\&=\frac{2^{a/2}}{\sqrt{\pi}}\Gamma\left(\frac{a+1}{2}\right)\left(\psi\left(\frac{a+1}{2}\right)+\log 2-c_1\right)\\&=\nu_a\kappa_a,\end{aligned}$$

and, using $\frac{\Gamma''(z)}{\Gamma(z)}=\psi'(z)+(\psi(z))^2$ (with primes denoting derivatives),

$$\begin{aligned}\mathrm{Cov}((\log X^2-c_1)^2,|X^a|)&=2\int_0^\infty\left[\left(\log x^2-c_1\right)^2-c_2\right]x^a\left(\tfrac{1}{\sqrt{2\pi}}\right)e^{-x^2/2}dx\\&=\frac{2^{a/2}}{\sqrt{\pi}}\int_0^\infty\left[\left(\log z+\log 2-c_1\right)^2-c_2\right]z^{(a-1)/2}e^{-z}dz\\&=\frac{2^{a/2}}{\sqrt{\pi}}\Gamma\left(\frac{a+1}{2}\right)\left(\frac{\Gamma''\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a+1}{2}\right)}+(\log 2-c_1)^2-c_2\right.\\&\qquad\left.+2(\log 2-c_1)\psi\left(\frac{a+1}{2}\right)\right)\\&=\nu_a\xi_a.\end{aligned}$$

Lemma 3 *Let X_1, X_2, and X_3 be jointly normal with $\mu_i=EX_i$ and $\sigma_{ij}=\mathrm{Cov}(X_i,X_j)$. Then*

$$\mathrm{Cov}(X_1,\exp X_3)=\sigma_{13}\exp\left(\mu_3+\tfrac{1}{2}\sigma_{33}\right) \tag{15}$$

and

$$\mathrm{Cov}(X_1X_2, \exp X_3) = (\sigma_{13}\sigma_{23} + \mu_1\sigma_{23} + \mu_2\sigma_{13}) \exp\left(\mu_3 + \tfrac{1}{2}\sigma_{33}\right). \tag{16}$$

Proof. Assume first that $\mu_i = 0$ and $\sigma_{ii} > 0$ for all i. Let $\mu_{i|j} = \sigma_{ij}\sigma_{jj}^{-1}X_j$ be the conditional mean of X_i, given X_j, and $\sigma_{ij|k} = \sigma_{ij} - \sigma_{ik}\sigma_{jk}\sigma_{kk}^{-1}$ the conditional covariance between X_i and X_j, given X_k. Then,

$$\begin{aligned}
\mathrm{Cov}(X_1, \exp X_3) &= \mathrm{Cov}(E(X_1|X_3), \exp X_3) \\
&= \sigma_{13}\sigma_{33}^{-1}\,\mathrm{Cov}(X_3, \exp X_3) \\
&= \sigma_{13}\exp\left(\tfrac{1}{2}\sigma_{33}\right), \\
\mathrm{Cov}(X_1X_2, \exp X_3) &= \mathrm{Cov}(E(X_1X_2|X_3), \exp X_3) \\
&= \mathrm{Cov}(\mu_{1|3}\mu_{2|3} + \sigma_{12|3}, \exp X_3) \\
&= \sigma_{13}\sigma_{23}\sigma_{33}^{-2}\,\mathrm{Cov}(X_3^2, \exp X_3) \\
&= \sigma_{13}\sigma_{23}\exp\left(\tfrac{1}{2}\sigma_{33}\right),
\end{aligned}$$

using the fact that, for a standard normal variate X,

$$\begin{aligned}
\mathrm{Cov}(X, \exp(bX)) &= \int_{-\infty}^{\infty} xe^{bx}\left(\tfrac{1}{\sqrt{2\pi}}\right)e^{-x^2/2}dx \\
&= \int_{-\infty}^{\infty} x\left(\tfrac{1}{\sqrt{2\pi}}\right)e^{-(x-b)^2/2+b^2/2}dx \\
&= b\exp\left(\tfrac{1}{2}b^2\right)
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Cov}(X^2, \exp(bX)) &= \int_{-\infty}^{\infty} (x^2-1)e^{bx}\left(\tfrac{1}{\sqrt{2\pi}}\right)e^{-x^2/2}dx \\
&= \int_{-\infty}^{\infty} (x^2-1)\left(\tfrac{1}{\sqrt{2\pi}}\right)e^{-(x-b)^2/2+b^2/2}dx \\
&= b^2\exp\left(\tfrac{1}{2}b^2\right).
\end{aligned}$$

The extension to the case where $\mu_i \neq 0$ for some i is straightforward, and any degenerate case follows upon taking the appropriate limit in the non-degenerate case.

Proof of Theorem 1. Write $z_t = k_t + w_t$, where $k_t = h_t - \mu$ and $w_t = \log u_t^2 - c_1$. Then, w_t and k_t have zero mean and are independent, and, for any integers i, j, l, we have $\mathrm{Cov}(k_t, k_{t-i}) = \phi^{|i|}\sigma^2$, $\mathrm{Cov}(k_t k_{t-i}, k_{t-j}) = 0$, and $\mathrm{Cov}(k_t k_{t-i}, k_{t-j} k_{t-l}) = (\phi^{|j|+|i-l|} + \phi^{|l|+|i-j|})\sigma^4$. Using these properties and the equalities

$$\sum_{l=-\infty}^{\infty} \phi^{|l|} = \frac{1+\phi}{1-\phi},$$

$$\sum_{l=-\infty}^{\infty} \phi^{|i+l|+|j+l|} = \phi^{|i-j|}\left(|i-j| + \frac{1+\phi^2}{1-\phi^2}\right),$$

we obtain, for $i, j \geq 0$,

$$\begin{aligned}
V(z_t, z_t) &= \sum_{l=-\infty}^{\infty} [\mathrm{Cov}(k_t, k_{t-l}) + \mathrm{Cov}(w_t, w_{t-l})] \\
&= \sum_{l=-\infty}^{\infty} \left(\phi^{|l|}\sigma^2 + I_{(l=0)} c_2\right) \\
&= \frac{1+\phi}{1-\phi}\sigma^2 + c_2, \\
V(z_t, z_t z_{t-j}) &= \sum_{l=-\infty}^{\infty} \mathrm{Cov}(w_t, w_{t-l} w_{t-j-l}) \\
&= \sum_{l=-\infty}^{\infty} I_{(l=j=0)} c_3 \\
&= I_{(j=0)} c_3,
\end{aligned}$$

and

$$
\begin{aligned}
V(z_t z_{t-i}, z_t z_{t-j}) &= \sum_{l=-\infty}^{\infty} [\mathrm{Cov}(k_t k_{t-i}, k_{t-l} k_{t-j-l}) + \mathrm{Cov}(w_t k_{t-i}, k_{t-l} w_{t-j-l}) \\
&\quad + \mathrm{Cov}(w_t k_{t-i}, w_{t-l} k_{t-j-l}) + \mathrm{Cov}(k_t w_{t-i}, k_{t-l} w_{t-j-l}) \\
&\quad + \mathrm{Cov}(k_t w_{t-i}, w_{t-l} k_{t-j-l}) + \mathrm{Cov}(w_t w_{t-i}, w_{t-l} w_{t-j-l})] \\
&= \sum_{l=-\infty}^{\infty} \Big[\left(\phi^{|l|+|j-i+l|} + \phi^{|j+l|+|l-i|} \right) \sigma^4 + I_{(l=-j)} \phi^{|i+j|} c_2 \sigma^2 \\
&\quad + I_{(l=0)} \phi^{|i-j|} c_2 \sigma^2 + I_{(l=i-j)} \phi^{|i-j|} c_2 \sigma^2 + I_{(l=i)} \phi^{|i+j|} c_2 \sigma^2 \\
&\quad + I_{(i=j\neq 0)} I_{(l=0)} c_2^2 + I_{(i=j=0)} I_{(l=0)} \left(c_4 - c_2^2 \right) \Big] \\
&= \left(|i-j| \, \phi^{|i-j|} + |i+j| \, \phi^{|i+j|} + \left(\phi^{|i-j|} + \phi^{|i+j|} \right) \frac{1+\phi^2}{1-\phi^2} \right) \sigma^4 \\
&\quad + 2 \left(\phi^{|i-j|} + \phi^{|i+j|} \right) c_2 \sigma^2 + I_{(i=j\neq 0)} c_2^2 + I_{(i=j=0)} (c_4 - c_2^2),
\end{aligned}
$$

giving (6)–(8). To establish (9), recall the definition of $Y^{i_1,\ldots,i_p}_{t_1,\ldots,t_p}$, from which

$$
\begin{aligned}
&\mathrm{Cov}\left(Y^{i_1,\ldots,i_p}_{t_1,\ldots,t_p}, Y^{i_{p+1},\ldots,i_{p+q}}_{t_{p+1}-l,\ldots,t_{p+q}-l} \right) \\
&\quad = E\left(Y^{i_1,\ldots,i_p}_{t_1,\ldots,t_p} Y^{i_{p+1},\ldots,i_{p+q}}_{t_{p+1}-l,\ldots,t_{p+q}-l} \right) - 1 \\
&\quad = E \exp\left(-\delta^{i_1,\ldots,i_p}_{t_1,\ldots,t_p} - \delta^{i_{p+1},\ldots,i_{p+q}}_{t_{p+1},\ldots,t_{p+q}} + \frac{1}{2} \sum_{j=1}^{p} i_j h_{t_j} + \frac{1}{2} \sum_{j=p+1}^{p+q} i_j h_{t_j-l} \right) \\
&\qquad \times E\left(\prod_{j=1}^{p} \left| \nu_{i_j}^{-1} u_{t_j}^{i_j} \right| \prod_{j=p+1}^{p+q} \left| \nu_{i_j}^{-1} u_{t_j-l}^{i_j} \right| \right) - 1.
\end{aligned}
$$

Now,

$$
\begin{aligned}
&E \exp\left(\frac{1}{2} \sum_{j=1}^{p} i_j h_{t_j} + \frac{1}{2} \sum_{j=p+1}^{p+q} i_j h_{t_j-l} \right) \\
&\qquad = \exp\left(\delta^{i_1,\ldots,i_p}_{t_1,\ldots,t_p} + \delta^{i_{p+1},\ldots,i_{p+q}}_{t_{p+1},\ldots,t_{p+q}} + \frac{\sigma^2}{4} \sum_{j=1}^{p} \sum_{j'=p+1}^{p+q} i_j i_{j'} \phi^{|t_j - t_{j'} + l|} \right),
\end{aligned}
$$

so

$$
\begin{aligned}
\mathrm{Cov}\left(Y^{i_1,\ldots,i_p}_{t_1,\ldots,t_p}, Y^{i_{p+1},\ldots,i_{p+q}}_{t_{p+1}-l,\ldots,t_{p+q}-l} \right) &= (B_l+1)(C_l+1) - 1 \\
&= B_l + (B_l+1) C_l.
\end{aligned}
$$

Summing over l gives (9), because $C_l = 0$ whenever $l \notin L$. Furthermore, by (13) and (15),

$$\begin{aligned}
&\mathrm{Cov}\left(z_t, Y^{i_1,\ldots,i_p}_{t_1-l,\ldots,t_p-l}\right) \\
&\quad = E\left(k_t \exp\left(-\delta^{i_1,\ldots,i_p}_{t_1,\ldots,t_p} + \frac{1}{2}\sum_{j=1}^{p} i_j h_{t_j-l}\right)\right) \\
&\qquad + \left(E\exp\left(-\delta^{i_1,\ldots,i_p}_{t_1,\ldots,t_p} + \frac{1}{2}\sum_{j=1}^{p} i_j h_{t_j-l}\right)\right) E\left(w_t \prod_{j=1}^{p}\left|\nu_{i_j}^{-1} u^{i_j}_{t_j-l}\right|\right) \\
&\quad = \frac{\sigma^2}{2}\sum_{j=1}^{p} i_j \phi^{|t-t_j+l|} + \sum_{j=1}^{p} I_{(t=t_j-l)}\kappa_{i_j},
\end{aligned}$$

which, upon summing over l, gives (10). Finally,

$$\begin{aligned}
\mathrm{Cov}\left(z_t z_{t-i}, Y^{i_1,\ldots,i_p}_{t_1-l,\ldots,t_p-l}\right) &= \mathrm{Cov}\Bigg(k_t k_{t-i} + k_t w_{t-i} + w_t k_{t-i} + w_t w_{t-i}, \\
&\qquad \exp\left(-\delta^{i_1,\ldots,i_p}_{t_1,\ldots,t_p} + \frac{1}{2}\sum_{j=1}^{p} i_j h_{t_j-l}\right) \prod_{j=1}^{p}\left|\nu_{i_j}^{-1} u^{i_j}_{t_j-l}\right|\Bigg) \\
&= T_1 + T_2 + T_3 + T_4,
\end{aligned}$$

say, with, using Lemma 2 and Lemma 3,

$$\begin{aligned}
T_1 &= \left(\frac{\sigma^2}{2}\sum_{j=1}^{p} i_j \phi^{|t-t_j+l|}\right)\left(\frac{\sigma^2}{2}\sum_{j=1}^{p} i_j \phi^{|t-i-t_j+l|}\right), \\
T_2 &= \left(\frac{\sigma^2}{2}\sum_{j=1}^{p} i_j \phi^{|t-t_j+l|}\right)\sum_{j=1}^{p} I_{(t-i=t_j-l)}\kappa_{i_j}, \\
T_3 &= \left(\frac{\sigma^2}{2}\sum_{j=1}^{p} i_j \phi^{|t-i-t_j+l|}\right)\sum_{j=1}^{p} I_{(t=t_j-l)}\kappa_{i_j}, \\
T_4 &= I_{(i=0)}\sum_{j=1}^{p} I_{(t=t_j-l)}\xi_{i_j} + I_{(i\neq 0)}\sum_{j,j'=1}^{p} I_{(t=t_j-l)} I_{(t-i=t_{j'}-l)}\kappa_{i_j}\kappa_{i_{j'}}.
\end{aligned}$$

Summing over l gives (11), which concludes the proof.

Proof of Lemma 1. Since $|\phi|^{|t_j - t_{j'} + l|} \le |\phi|^{-|t_j - t_{j'}| + |l|}$,

$$B_l \le \exp\left(\frac{\sigma^2}{4}\sum_{j=1}^{p}\sum_{j'=p+1}^{p+q} i_j i_{j'} |\phi|^{|t_j - t_{j'} + l|}\right) - 1 \le \exp(a|\phi|^{|l|}) - 1.$$

By an argument of symmetry,

$$\exp(-a|\phi|^{|l|}) - 1 \le B_l \le \exp(a|\phi|^{|l|}) - 1$$

and so, because $\exp(z) - 1 \ge 1 - \exp(-z)$ for any z,

$$|B_l| \le \exp(a|\phi|^{|l|}) - 1.$$

Therefore,

$$\begin{aligned} |B - B(I)| &\le \sum_{l=I+1}^{\infty}(|B_{-l}| + |B_l|) \le 2\sum_{l=I}^{\infty}\left(\exp\left(a|\phi|^l\right) - 1\right) \\ &= 2\sum_{l=I}^{\infty}\sum_{k=1}^{\infty}\frac{a^k|\phi|^{kl}}{k!} = 2\sum_{k=1}^{\infty}\left(\frac{a^k}{k!}\right)\left(\frac{|\phi|^{kI}}{1-|\phi|^k}\right) \\ &< \frac{2}{1-|\phi|}\sum_{k=1}^{\infty}\frac{\left(a|\phi|^I\right)^k}{k!} = 2\frac{\exp(a|\phi|^I) - 1}{1-|\phi|}. \end{aligned}$$

References

[1] Andersen, T.G., H.-J. Chung, and B.E. Sørensen, 1999, Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study, Journal of Econometrics 91, 61–87.

[2] Andersen, T.G. and B.E. Sørensen, 1996, GMM estimation of a stochastic volatility model: A Monte Carlo study, Journal of Business & Economic Statistics 14, 328–352.

[3] Andersen, T.G. and B.E. Sørensen, 1997, GMM and QML ssymptotic standard deviations in stochastic volatility models: Comments on Ruiz (1994), Journal of Econometrics 76, 397–403.

[4] Breusch, T, H.L. Qian, P. Schmidt, and D. Wyhowski, 1999, Redundancy of moment conditions, Journal of Econometrics 91, 89–111.

[5] Chib, S., F. Nardari, and N. Shephard, 2002, Markov chain Monte Carlo methods for stochastic volatility models, Journal of Econometrics 108, 281–316.

[6] Clark, P.K., 1973, A subordinated stochastic process model with finite variance for speculative prices, Econometrica 41, 135–155.

[7] Danielsson, J., 1994, Stochastic volatility in asset prices: Estimation with simulated maximum likelihood, Journal of Econometrics 54, 375–400.

[8] Danielsson, J. and J.F. Richard, 1993, Accelerated Gaussian importance sampler with application to dynamic latent variable models, Journal of Applied Econometrics 8 (Suppl.), S153–S173.

[9] Duffie, D. and K.J. Singleton, 1993, Simulated moments estimation of Markov models of asset prices, Econometrica 61, 929–952.

[10] Fedorov, V.V., 1972, Theory of optimal experiments (Academic Press, New York).

[11] Fridman, M. and L. Harris, 1998, A maximum likelihood approach for non-Gaussian stochastic volatility models, Journal of Business & Economic Statistics 16, 284–291.

[12] Gallant, A.R., D.A. Hsieh, and G.E. Tauchen, 1997, Estimation of stochastic volatility models with diagnostics, Journal of Econometrics 81, 159–192.

[13] Ghysels, E., A.C. Harvey, and E. Renault, 1996, Stochastic volatility, in: G.S. Maddala, ed., Handbook of statistics: Statistical methods in finance, Vol. 14 (North-Holland, Amsterdam) 119–191.

[14] Gouriéroux, C., A. Monfort, and E. Renault, 1993, Indirect inference, Journal of Applied Econometrics 8, S85–S118.

[15] Hansen, L.P., 1982, Large sample properties of generalized methods of moments estimators, Econometrica 50, 1029–1054.

[16] Hansen, L.P., J. Heaton, and A. Yaron, 1996, Finite-sample properties of some alternative GMM estimators, Journal of Business & Economic Statistics 14, 262–280.

[17] Harvey, A.C, E. Ruiz, and N. Shephard, 1994, Multivariate stochastic volatility models, The Review of Economic Studies 61, 247–264.

[18] Hull, J. and A. White, 1987, The pricing of options on assets with stochastic volatilities, The Journal of Finance 17, 281–300.

[19] Jacquier, E., N.G. Polson, and P.E. Rossi, 1994, Bayesian analysis of stochastic volatility models, Journal of Business & Economic Statistics 12, 69–87.

[20] Kim, S., N. Shephard, and S. Chib, 1998, Stochastic volatility: optimal likelihood inference and comparison with ARCH models, The Review of Economic Studies 45, 361–393.

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

[22] Monfardini, C., 1998, Estimating stochastic volatility models through indirect inference, The Econometrics Journal 1, C113–C128.

[23] Nelson, D.B., 1988, Time series behavior of stock market volatility and returns, Ph.D. dissertation, M.I.T..

[24] Newey, W.K. and R.J. Smith, 2000, Asymptotic bias and equivalence of GMM and GEL estimators, Discussion paper No. 01/517, University of Bristol.

[25] Newey, W.K. and R.J. Smith, 2001, Higher order properties of GMM and generalized empirical likelihood estimators, Discussion paper, M.I.T..

[26] Ruiz, E., 1994, Quasi-maximum likelihood estimation of stochastic volatility models, Journal of Econometrics 63, 289–306.

[27] Sandmann, G. and S.J. Koopman, 1998, Estimation of stochastic volatility models via Monte Carlo maximum likelihood, Journal of Econometrics 87, 271–301.

[28] Shephard, N., 1996, Statistical aspects of ARCH and stochastic volatility, in: D.R. Cox, O.E. Barndorff-Nielsen, and D.V. Hinkley, eds., Time series models in econometrics, finance and other fields (Chapman & Hall, London) 1–67.

[29] Tauchen, G.E. and M. Pitts, 1983, The price variability-volume relationship on speculative markets, Econometrica 51, 484–505.

[30] Taylor, S., 1986, Modelling financial time series (Wiley, Chichester).

[31] Wright, J.H., 1999, A new estimator of the fractionally integrated stochastic volatility model, Economics Letters 63, 295–303.