ORIGINAL PAPER

Iterative QML estimation for asymmetric stochastic volatility models

Paolo Chirico[1]

Abstract
The paper illustrates a new procedure for estimating asymmetric stochastic volatility models. These models shape the asymmetric effect of negative and positive financial returns on the expected volatility, behaviour often observed in the stock prices, and known as "leverage effect". The procedure is based on the iterative application of the quasi-maximum likelihood (QML) method and is proposed as an alternative to the procedure presented by Harvey and Shephard in 1996 and based on the application of the QML method on a modified auxiliary model. The estimation results generally converge to constant values after a few iterations. The volatility predictor provided by the new method is conceptually similar to the EGARCH predictor and different from the predictor of the other procedure. A simulation study shows that the iterative QML method provides parameter estimators with RMSEs decreasing as series length increases. The distribution of the estimates is approximately normal, and the approximation improves as the series size increases. Empirical applications of the method provide results similar to ones of the method known in literature. However, the two methods provide two different predictors and smoothers of volatility, which should be compared on a case-by-case basis.

Keywords Asymmetric stochastic volatility · Leverage effect · Iterative quasi maximum likelihood

JEL Classification C13 · C32 · C58

1 Introduction

According with a very common pattern for financial returns, r_t, the term volatility refers to the coefficient σ_t into the equation:

✉ Paolo Chirico
paolo.chirico@uniupo.it

1 Department of Law and Political, Economic and Social Sciences, Università del Piemonte Orientale, Alessandria, Italy

$$r_t = \mu_t + \sigma_t \varepsilon_t \quad \varepsilon_t \sim NID(0,1) \tag{1}$$

where ε_t is a unit-variance innovation, Gaussian distributed in the simplest case; when σ_t does not depend on ϵ_t, μ_t is the expected value of r_t given all information known at time $t-1$ (Taylor 1986).

Two main classes of volatility models are known in literature:

- Models for conditional heteroskedasticity;
- Stochastic Volatility models.

In the models for conditional heteroskedasticity, generally identified as GARCH models (Bollerslev 1986), volatility is a function of the previous information, I_{t-1}, and σ_t^2 corresponds to the conditional variance of the returns: $\sigma_t^2 = Var(r_t \mid I_{t-1})$.

An alternative approach in volatility modelling consists in considering σ_t^2 as a latent stochastic process, whose logarithm, h_t, is usually represented as an autoregressive process, in most cases, of order one:

$$h_{t+1} = \omega + \beta h_t + \eta_t \quad \eta_t \sim NID(0, \sigma_\eta^2) \tag{2}$$

Models in the form (1)–(2) are reported in literature with the general name of *Stochastic Volatility* (SV) models (Melino and Turnbull 1990; Taylor 1994) and are discrete approximations to various diffusion processes proposed in the asset-pricing theory (Hull and White 1987; Wiggins 1987).

Unlike GARCH family models, which are easily estimated by the *Maximum Likelihood* (ML) method, the ML estimation of SV models is a tough challenge since the likelihood is defined by a T-dimensional integral, which is hard to manage (Harvey and Shephard 1996).

In accord with Sandmann and Koopman (1998), methods to face SV models estimation can be subdivided into two groups: (i) methods oriented to rebuild the exact likelihood of the SV model or a related model; (ii) methods based on more workable, but sub-optimal methods. Among all the methods, the Quasi Maximum Likelihood (QML) method (Ruiz 1994; Harvey et al. 1994) appears to be a good compromise between simplicity and efficiency: it maximizes a function that is not the actual exact likelihood function, but it provides an optimal linear estimator (and predictor) of h_t, consistent and asymptotically Gaussian, according with the results of Dunsmuir (1979).

The method is based on the Kalman filter (Harvey 1990), which filters past and present volatilities and predicts future ones; Ruiz (1994) suggest that the method works well for sample sizes usually used in financial economics.

On the other hand, the application of the basic[1] QML faces some issues when the volatility innovation, η_t, and the return innovation, ε_t, are correlated, as explained in Sect. 2. This correlation is fundamental to model the asymmetric effect of negative

---

[1] In the paper, the term "basic" is used to identify the method described by Ruiz (1994) and Harvey et al. (1994).

and positive returns on the expected volatility, behaviour often observed in stock prices, and known as *leverage effect* (Christie 1982; Engle and Ng 1993).

This paper presents a new procedure for applying the QML method to asymmetric SV models, which produces a volatility predictor different from the one known in the literature, preferable to the latter in some cases. A brief presentation of the known procedure is reported in Sect. 2, while the new procedure is illustrated in Sect. 3. In Sect. 4, a simulation study evaluates the properties of the estimators obtained with the new proposal. Section 5 illustrates the application of the old and new procedure on three financial series. The achieved goals are summarized in Conclusions.

2 Leverage effect and stochastic volatility

In financial series increases of volatility are often observed after negative returns, and generally the greater the extent of the loss the greater the increase. This evidence is called "leverage effect" as price drop in a stock decreases the value of the firm equity, and increases the leverage-ratio. The increased leverage-ratio will involve higher risk on the equity which will be more volatile during next period (Black 1976).

A way to include the leverage effect in (2) consists in assuming the volatility innovation η_t correlated with the return innovation ε_t, as in the following general Gaussian SV model:

$$\begin{aligned} x_t &= \exp(h_t/2)\varepsilon_t \\ h_{t+1} &= \omega + \beta h_t + \eta_t \end{aligned} \qquad \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim NID\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \gamma \\ \gamma & \sigma_\eta^2 \end{bmatrix} \right) \qquad (3)$$

where: $x_t = r_t - \mu_t$ is the mean-adjusted return on an asset, simply "return" in the following; $h_t = \ln \sigma_t^2$ is the log-square volatility; ε_t and η_t are zero-mean Gaussian innovations, serially independent, with $E(\eta_t \varepsilon_{t-k}) = \gamma$ if $k = 0$, and zero otherwise; then, the correlation between ε_t and η_t is $\rho = \gamma/\sigma_\eta$.

In Model (3), the leverage effect on h_{t+1} corresponds to:

$$E(\eta_t \mid \varepsilon_t) = \gamma \varepsilon_t \qquad (4)$$

which is proportional to the size of ε_t and of opposite sign if $\gamma < 0$.

The basic QML estimation of a Gaussian SV model would consists in the estimation, via Kalman Filter, of the auxiliary state space model:

$$\begin{aligned} y_t &= -1.27 + h_t + \xi_t \\ h_{t+1} &= \omega + \beta h_t + \eta_t \end{aligned} \qquad \begin{bmatrix} \xi_t \\ \eta_t \end{bmatrix} \sim ID\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4.93 & 0 \\ 0 & \sigma_\eta^2 \end{bmatrix} \right) \qquad (5)$$

where: $y_t = \ln x_t^2$, and $\xi_t = \ln \varepsilon_t^2 + 1.27$; -1.27 and 4.93 are the values of $E(\ln \varepsilon_t^2)$ and $Var(\ln \varepsilon_t^2)$ respectively, when ε_t is Gaussian (Zelen and Severo 1972).

Unfortunately, the basic QML does not allow to estimate the parameter γ (or ρ) directly since the disturbances ξ_t and η_t are uncorrelated when the distribution of

ε_t and η_t is symmetric, e.g. Gaussian or Student's t, so that the original covariance (correlation) between ε_t and η_t is definitely lost in (5) (Harvey et al. 1994).

Harvey and Shephard (1996) proposed a method, here briefly named HS-QML, to estimate the parameters of a SV model with correlated disturbances. According with this method, the parameters of (3) can be estimated applying the QML method to the time-varying linear state space model[2]:

$$
\begin{aligned}
y_t &= -1.27 + h_t + \xi_t \\
h_{t+1} &= \omega + 0.80\gamma s_t + \beta h_t + \eta_t^*
\end{aligned}
$$

$$
\begin{bmatrix} \xi_t \\ \eta_t^* \end{bmatrix} \mid s_t \sim ID\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4.93 & 1.11\gamma s_t \\ 1.11\gamma s_t & \sigma_\eta^2 - 0.64\gamma^2 \end{bmatrix} \right) \tag{6}
$$

where s_t indicates the sign of x_t: it is the equal to 1 (-1) when x_t, i.e. ε_t, is positive (negative).

The one step ahead predictor of h_t provided by the Kalman Filter (see Appendix A) combines the mechanisms of the EGARCH (Nelson 1991) and the Threshold-ARCH predictors (Glosten et al. 1993; Zakoian 1994):

$$
\hat{h}_{t+1|t} = \omega + \beta \hat{h}_{t|t-1} + 0.80\gamma s_t + \kappa_t(y_t - \hat{y}_{t|t-1}) \tag{7}
$$

$$
\kappa_t = \frac{\beta P_{t|t-1} + 1.11\gamma s_t}{P_{t|t-1} + 4.93} \tag{8}
$$

where κ_t is the gain of the Kalman filter (Harvey 1990); $\hat{y}_{t|t-1} = -1.27 + \hat{h}_{t|t-1}$ is the one step ahead prediction of y_t; $P_{t|t-1} = E[(\hat{h}_{t|t-1} - h_t)^2]$ is the MSE of $\hat{h}_{t|t-1}$.

If $\gamma < 0$, Formula (8) entails that: $(\kappa_t \mid P_{t|t-1}, s_t = -1) > (\kappa_t \mid P_{t|t-1}, s_t = +1)$. Therefore the leverage effect predicted by (7) can be broken down into two parts: (i) a part, γs_t, depending on the sign but not on the size of ε_t (as in the Threshold-ARCH model); (ii) a part, $\kappa_t(y_t - \hat{y}_{t|t-1})$, depending on the sign and the size of ε_t .[3] This forecast only partially reflects the characteristics of the leverage effect as expressed in (4).

3 Iterative QML for asymmetric SV models

Having assumed that η_t, and ε_t are bivariate normal with $E(\eta_t \varepsilon_t) = \gamma$, the auxiliary model (5) can be rewritten as:

---

[2] See "Appendix A".

[3] $(y_t - \hat{y}_{t|t-1})$ can be viewed as an approximation of $\xi_t = \ln \varepsilon_t^2 + 1.27 = 2\ln|\varepsilon_t| + 1.27$.

$$y_t = -1.27 + h_t + \xi_t$$
$$h_{t+1} = \omega + \gamma\varepsilon_t + \beta h_t + \eta_t^+$$
$$\begin{bmatrix} \xi_t \\ \eta_t^+ \end{bmatrix} \sim ID\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4.93 & 0 \\ 0 & \sigma_\eta^2 - \gamma^2 \end{bmatrix}\right) \tag{9}$$

where the disturbance η_t^+ represents the exogenous innovation on h_t.

As $\varepsilon_t = x_t / \exp(h_t/2)$, Model (9) is not linear, then it cannot be estimated by means of basic QML. Nevertheless, we can perform the following procedure:

1. ε_t is initially set to x_t/s_X, being s_X the sample standard deviation of the series of returns $(x_1, x_2, \ldots, x_T)$;
2. Model (9), now linear, is estimated by basic QML;
3. ε_t is updated by $x_t / \exp(\tilde{h}_t/2)$, where $\tilde{h}_t$ are the smoothed h_t, provided by the Kalman smoother;
4. steps 2 and 3 are repeated successively according with a pre-set stopping rule[4].

Empirical evidence shows that the parameter estimates converge to realistic values after few steps. The trick of the iterative procedure, IQML in the following, consists in smoothing ε_t and h_t not conjointly: treating separately ε_t and h_t does not compromise the linearity of the second equation into (9).

The IQML can be viewed like a variant of the EM approach (Dempster et al. 1977) proposed by Shumway and Stoffer (1982) for smoothing and forecasting time series, but some differences should be highlighted. In the case of Model (9), the approach of Shumway and Stoffer (1982) would consist in: (i) setting initial values of the model parameters; (ii) smoothing h_t in order to build a likelihood function (expectation step); (iii) estimating the model parameters maximizing the likelihood function (maximization step); (iv) iterating steps (ii)–(iii) until the estimates and the likelihood function are stable. Nevertheless the direct smoothing of h_t (step ii) is complicated by the non-linearity of the model. This problem is bypassed in IQML because ε_t and h_t are smoothed separately: ε_t is smoothed before the maximisation (estimation) step, whereas h_t is smoothed after the maximization step in order to update ε_t. As a result the maximisation step differs between the algorithms: in IQML it consists in maximizing a pseudo (quasi) log-likelihood given y_t and $\tilde{\varepsilon}_t$; in EM it consists in the ML estimation of a regression model, the second equation in (9), given $\tilde{h}_t$.[5] The IQML procedure is more viable than the EM approach described above, but involves an inevitable loss of efficiency due to smoothing ε_t and h_t separately.

It is interesting to note the form of the predictor of h_t provided by the Kalman Filter:

---

[4] In simulation study (Sect. 4) and empirical applications (Sect. 5), the procedure is stopped when the increment of the quasi log-likelihood becomes less than 0.001. With this stopping rule, the relative difference in the estimates between the last two iterations becomes less than 0.1%

[5] Given $\tilde{h}_t$ (and $\tilde{\varepsilon}_t = x_t / \exp(\tilde{h}_t/2)$), the model estimation concerns only the second equation in (9).

$$\hat{h}_{t+1|t} = \omega + \beta \hat{h}_{t|t-1} + \gamma \tilde{\varepsilon}_t + \kappa_t(y_t - \hat{y}_{t|t-1}) \tag{10}$$

where κ_t is the gain of the Kalman Filter. Since Model (9) satisfies the steady-state conditions,[6] κ_t converges to a constant that we name α. Therefore, the steady-state predictor of h_t can be formalized as:

$$\hat{h}_{t+1|t} = \omega + \beta \hat{h}_{t|t-1} + \gamma \tilde{\varepsilon}_t + \alpha \xi(|\tilde{\varepsilon}_t|) \tag{11}$$

where the function $\xi(|\tilde{\varepsilon}_t|) = 2\ln|\tilde{\varepsilon}_t| + 1.27$ is a monotonically increasing, mean-corrected, function of the (approximated) magnitude of ε_t. Predictor (11) shows a clear similarity with the EGARCH predictor in which $\xi(|\varepsilon_t|) = |\varepsilon_t| - \sqrt{2/\pi}$ is the mean-corrected magnitude of ε_t when $\varepsilon_t \sim N(0,1)$. Predictor (11) may be also viewed as a Log-GARCH predictor (Geweke 1986; Pantula 1986) with leverage effect.

The leverage effect predicted by (11) tries to replicate the form of the leverage effect Model (4): it is proportional to the amplitude of the estimated return innovation, $\tilde{\varepsilon}_t$, and of opposite sign (if $\gamma < 0$). Formulas (7) and (11) show that the HS-QML and IQML methods involve two different predictors of h_t, whose performances are evaluated in Sects. 4 and 5.

3.1 Student's t return innovations

The IQML method can be generalized to the case where ε_t has a scaled Student's t-distribution with v degrees of freedom, scaled in order to have unit variance, i.e. $\varepsilon_t \sim t_v \sqrt{(v-2)/v}$. In this case (see "Appendix B"):

$$E[\ln \varepsilon_t^2] = g_1(v) = -1.27 + \ln(v/2 - 1) - \psi_0(v/2) \tag{12}$$

$$Var[\ln \varepsilon_t^2] = g_2(v) = 4.93 + \psi_1(v/2) \tag{13}$$

being ψ_0 and ψ_1 the *digamma* and *trigamma* function, respectively (Davis 1972).

If the parameter v is assumed known, the model estimation procedure is the IQML described above, with the values -1.27 and 4.93 now replaced by the values of $g_1(v)$ and $g_2(v)$, respectively. Alternatively, -1.27 and 4.93 are replaced by the parametric formulas of $g_1(v)$ and $g_2(v)$, and v is treated as an additional unknown parameter.

---

[6] We assume $0 < \beta < 1$ and $\sigma_\eta^2 > \gamma^2$ (i.e. $-1 < \rho < 1$). For more details about the convergence properties of the Kalman Filter see Harvey (1990) and Hamilton (1994).

Table 1 Simulations results of the HS-QML method ($n = 1000$)

Parameters	Actual Values	Average Estimates					
		$T = 1000$		$T = 3000$		$T = 6000$	
β	*0.975*	0.948	(0.088)	0.970	(0.028)	0.973	(0.009)
$\ln \sigma_\eta^2$	*− 4.605*	− 4.437	(1.185)	− 4.590	(0.692)	− 4.604	(0.432)
ρ	*− 0.300*	− 0.335	(0.298)	− 0.304	(0.156)	− 0.303	(0.101)
β	*0.975*	0.968	(0.034)	0.974	(0.007)	0.975	(0.005)
$\ln \sigma_\eta^2$	*− 4.605*	− 4.596	(0.708)	− 4.617	(0.353)	− 4.612	(0.249)
ρ	*− 0.900*	− 0.909	(0.132)	− 0.911	(0.079)	− 0.907	(0.058)

Italic values are the actual parameter values used to simulate the 1000 series

Table 2 Simulations results of the IQML method ($n = 1000$)

Parameters	Actual Values	Average estimates					
		$T = 1000$		$T = 3000$		$T = 6000$	
β	*0.975*	0.945	(0.077)	0.969	(0.019)	0.972	(0.009)
$\ln \sigma_\eta^2$	*− 4.605*	− 4.373	(1.167)	− 4.590	(0.645)	− 4.599	(0.430)
ρ	*− 0.300*	− 0.313	(0.284)	− 0.309	(0.148)	− 0.298	(0.095)
β	*0.975*	0.967	(0.028)	0.974	(0.007)	0.975	(0.004)
$\ln \sigma_\eta^2$	*− 4.605*	− 4.438	(0.646)	− 4.612	(0.324)	− 4.612	(0.210)
ρ	*− 0.900*	− 0.875	(0.127)	− 0.903	(0.069)	− 0.900	(0.044)

Italic values are the actual parameter values used to simulate the 1000 series

4 Finite sample properties of the IQML estimator

By means of a simulation study, Harvey and Shephard (1996) provided empirical results in favour of the consistency of the estimates obtained with the HS-QML method. In the study, series of different lengths, T, were simulated 1000 times ($n = 1000$) from Model (3) with "empirically reasonable" parameters values. On each series, the model parameters were estimated using the HS-QML method, under some practical constraints.[7]

Table 1 reports some estimation results of that simulation study: the average and the root mean square error, RMSE, (figures in brackets) of the estimates of β, $\ln \sigma_\eta^2$ and ρ; $\ln \sigma_\eta^2$ was preferred to σ_η^2 because the estimates of the first are closer to be normally distributed then ones of the second. We can note that the RMSEs decrease as the series length increases and, ceteris paribus, they decrease if the absolute value of ρ (i.e. γ) increases.

The same approach, based on simulations, is followed here to asses the finite sample properties of the IQML estimators.[8] With the same settings adopted by Harvey and Shephard, 1000 series were simulated from Model (3), and on each

---

[7] See Harvey and Shephard (1996) for more details.

[8] The procedure was performed using the econometric software Gretl, version 2022a.

Table 3 IQML—p value of normality tests ($n = 1000$)

	$T = 3000$		$T = 6000$	
$\rho =$	-0.900	-0.300	-0.900	-0.300
Test for normality of $\hat{\omega}$				
Shapiro–Wilk test	0.244	0.293	0.952	0.090
Jarque–Bera test	0.109	0.968	0.723	0.088
Test for normality of $\hat{\gamma}$				
Shapiro–Wilk test	0.018	0.179	0.271	0.135
Jarque–Bera test	0.021	0.174	0.144	0.392
Test for normality of $\hat{\beta}$				
Shapiro–Wilk test	0.000	0.000	0.000	0.000
Jarque–Bera test	0.000	0.000	0.000	0.000
Test for normality of $\ln \hat{\sigma}_\eta^2$				
Shapiro–Wilk test	0.090	0.000	0.380	0.178
Jarque–Bera test	0.009	0.000	0.330	0.031
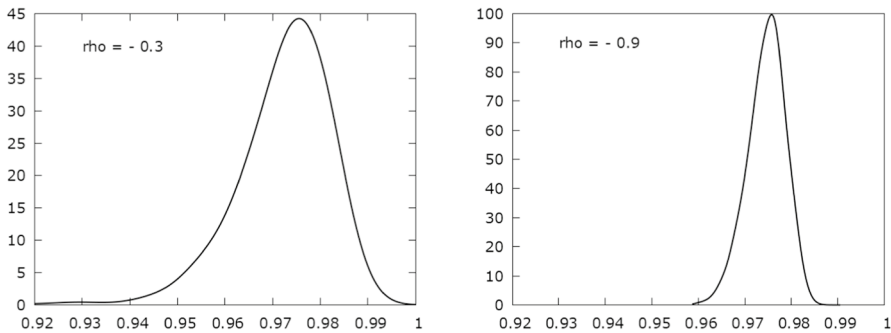
Fig. 1 Estimated density of $\hat{\beta}$ in the simulation study ($n = 1000$)

series the model was estimated with the IQML method (Table 2). As in HS-QML, the RMSEs of the IQML estimators decrease as the strength of the correlation increases and as the length of the series increases. The RMSEs of the new method are slightly lower than ones of the other method.

The normal distribution of the IQML estimates has been tested using the Shapiro-Wilks and the Jarque-Bera tests. Table 3 reports the p-values of the tests on the estimates of ω, γ, β, and $\ln \sigma_\eta^2$, both with $T = 3000$ and $T = 6000$. With series of length $T = 6000$, the estimates of ω, γ and $\ln \sigma_\eta^2$ can be already considered normally distributed with a significance level of 0.03. On the other hand, the sampling distribution of $\hat{\beta}$ cannot yet be approximated by the normal, this is because the simulated value, 0.975, is very close to the upper limit in case of stationary volatility ($\beta = 1$). As a result, the finite sample distribution of $\hat{\beta}$ has a slightly longer left tail than the right one, although the skewness decreases the
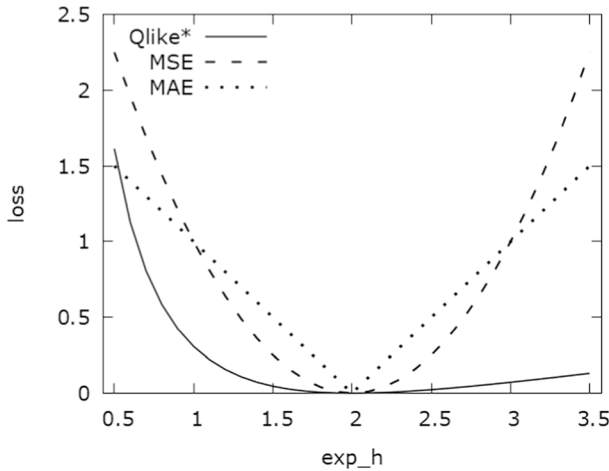
Fig. 2 Behaviour of the loss functions given $\sigma^2 = 2$

more negative the correlation ρ (Fig. 1). In this case the Gaussian approximation should be possible with series length greater than those considered.

4.1 Goodness of the IQML filtered and smoothed volatilities

A simulation was conducted to evaluate the goodness of the filtered and smoothed volatilities provided by the IQML and HS-QML methods. To this end: (i) 100 series of length $T = 1000$ were simulated from model (3) with $\beta = 0.97$, $\rho = -0.90$, $\ln(\sigma_\eta^2) = -4$ (i.e. $\gamma = -0.122$ and $\sigma_\eta^2 = 0.0183$); (ii) on each series, the IQML and HS-QML methods were applied to filter and smooth the log-square volatility h_t using the estimated parameters provided by each method, then filtered and smoothed σ_t^2s were obtained by exponential transformation of the corresponding log-square volatilities; (iii) the closeness of the filtered and smoothed σ_t^2s to the simulated ones was measured on each series using the (average) loss functions MSE, MAE and Qlike (Patton 2011; Hansen and Lunde 2005):

$$\text{MSE} = T^{-1} \sum_{t=1}^{T} (\sigma_t^2 - \exp h_t)^2 \tag{14}$$

$$\text{MAE} = T^{-1} \sum_{t=1}^{T} |\sigma_t^2 - \exp h_t| \tag{15}$$

$$\text{Qlike*} = T^{-1} \sum_{t=1}^{T} \left(\frac{\sigma_t^2}{\exp h_t} - 1 - \ln \frac{\sigma_t^2}{\exp h_t} \right) \tag{16}$$

Table 4 Overall average of the loss functions on the simulated series ($n = 100; T = 1000$)

Method	Filtered σ_t^2s			Smoothed σ_t^2s		
	MSE	MAE	Qlike*	MSE	MAE	Qlike*
HS-QML	1.000	1.000	1.000	1.000	1.000	1.000
IQML[1]	0.807	0.886	0.873	0.767	0.879	0.826
n^{*2}	87	91	86	93	91	89

Results as fraction of the corresponding HS-QML loss

$n^*=$ number of series where IQML determines the lowest loss

The loss function (16) corresponds to the variant of Qlike proposed by Patton (2011, p. 252) in order to make the function homogeneous. This variant is formally the relative difference of σ_t^2 from $\exp h_t$ minus the corresponding log-difference. Given the relationship between the two differences, Qlike* grows as the gap between σ_t^2 and $\exp h_t$ increases, and grows more when the gap is negative, i.e. σ_t^2 is underestimated (Fig. 2). On the other hand, MSE and MAE present symmetric effects of overestimation and underestimation, but MSE is more sensitive than MAE to large gaps between σ_t^2 and $\exp h_t$ (due to sharp changes in volatility or outliers).

Table 4 reports the overall average of the loss functions for the methods,[9] and the the number of series, n^*, where IQML determines the lowest loss. We can see that the values of all three loss functions are lower when the IQML method is used, both in case of filtered and in case of smoothed σ_t^2s. These results seem to suggest that the IQML method provides filtered and smoothed σ_t^2s that fit better the actual ones. In particular, the IQML method appears to better limit: (i) large gaps between actual volatility and filtered (smoothed) volatility; (ii) underestimation of actual volatility.

Finally, the n^* counter shows that the IQML method outperforms the HS-QML in a large percentage of simulated series.

5 Empirical Applications

The IQML method was applied on several financial series for the estimation of the asymmetric SV model (3). This section illustrates the application of the method on three series of financial indices:

- NASDAQ Composite (IXIC)
- DAX index (GDAXI)
- CAC40 index (FCHI)

---

[9] Results are reported as fraction of the corresponding HS-QML loss. Since the functions are homogeneous, the use of relative values is not affected by the scale of measurement of the returns (percentages or decimals).

Table 5 Comparison of the methods on three financial series

Method	Param	IXIC		GDAXI		FCHI	
		Estimate	SE	Estimate	SE	Estimate	SE
HS-QML	ω	0.025	0.010	0.012	0.006	0.014	0.006
	γ	− 0.157	0.046	− 0.264	0.042	− 0.270	0.049
	β	0.966	0.013	0.962	0.011	0.950	0.012
	σ_{η^+}	0.137	0.047	0.086	0.043	n.s.[1]	
	lly	− 1208.6		− 1151.2		− 1124.9	
IQML	ω	0.014	0.005	− 0.002	0.008	− 0.002	0.003
	γ	− 0.157	0.036	− 0.220	0.041	− 0.204	0.038
	β	0.977	0.007	0.956	0.015	0.973	0.007
	σ_{η^+}	0.072	0.038	0.137	0.058	n.s.[1]	
	lly	− 1206.3		− 1153.0		− 1123.4	

n.s. = non-significant

Table 6 MSE and Qlike losses of the filtered and smoothed σ_t^2

Index	Method	Filtered σ_t^2s		Smoothed σ_t^2s	
		MSE	Qlike*	MSE	Qlike*
IXIC	HS-QML	1.000	1.000	1.000	1.000
	IQML	0.994	0.990	1.012	1.016
GDAXI	HS-QML	1.000	1.000	1.000	1.000
	IQML	0.989	1.025	0.966	0.999
FCHI	HS-QML	1.000	1.000	1.000	1.000
	IQML	0.953	1.016	0.953	1.014

Results as fraction of the corresponding HS-QML loss

in the period from 04-01-2021 to 30-12-2022. The returns we consider are percentage log differences of the daily index closing values. The estimation results are compared with those obtained with the HS-QML method.[10]

Table 5 reports the estimates of ω, γ, β, and the standard deviation of the exogenous innovation on h_t (i.e. σ_{η^+}). The standard errors of the estimates are calculated on the basis of the "Outer Product of the Gradient" (OPG) method.[11] The quasi log-likelihood, lly, of models (6) and (9) is also reported.

The two methods provide fairly close estimates, also standard errors are pretty close. We note that the IQML standard errors are moderately smaller than the corresponding HS-QML standard errors in the IXIC and FCHI series, but not in GDAXI. The quasi-log likelihood values, lly, of the two methods are also very close in each

---

[10] The HS-QML method is applied using the equivalent model form (A.5) explained in Appendix (A). This form puts in evidence the standard deviation of the exogenous innovation on h_t.

[11] See Gretl User's Guide, version of January 2021, p. 222 (Cottrell and Lucchetti 2021).

series: the *lly* of IQML is slightly higher in IXIC and FCHI, and slightly lower in GDAXI.

The goodness of the filtered and smoothed σ_t^2s provided by the methods is assessed by the MSE (14) and Qlike* (16) loss functions computed using the square return x_t^2 as proxy of σ_t^2 (see Table 6). As stated by Patton (2011), these loss functions are robust to noise when the volatilities proxy are the square returns. Based on these results, neither method appears clearly better than the other: the IQML seems little better for the IXIC filtered σ_t^2s and the GDAXI smoothed σ_t^2s; the HS-QML seems little better for the IXIC smoothed σ_t^2s; for the FCHI series, IQML seems preferable if the criterion is MSE, but HS-QML could be better if Qlike is the criterion. Nevertheless, all differences in the loss functions are too small to highlight a clear superiority of one method over the other. In the case of financial series, the choice of the most appropriate method should take place on a case-by-case basis, considering more than one criterion.

6 Conclusion

The IQML method consists in iterating the basic QML method over an asymmetric SV model. The procedure is made possible using a proxy of the return innovation ε_t. This simple procedure allows the user to estimate the parameters of an SV model in which the return innovation and volatility innovation are correlated. This goal can also be achieved with the modified QML method proposed by Harvey and Shephard (1996) (HS-QML), but the two methods provide different volatility predictors. The IQML predictor is conceptually similar to the EGARCH predictor, whereas the HS-QML predictor is more similar to the Threshold-ARCH predictor. A simulation study shows that the IQML filtered and smoothed square volatilities generally fit simulated σ_t^2s better than those of the other model do. On the other hand, empirical applications suggest comparing the two methods, using loss functions, to identify the most suitable for the series under study.

The simulation study also shows that IQML estimators exhibit decreasing RMSEs as the series length increases and finite sample distributions that can be approximated by the Gaussian distribution; the approximation improves as the sample size increases.

The IQML method can be viewed as a variant of the Expectation-Maximization (EM) algorithm proposed by Shumway and Stoffer (1982), although the algorithms differ in some methodological aspects as specified in Sect. 3.

Finally, the method can be applied to the case with Student's *t* return innovations in order to treat returns with high kurtosis.

Appendix A Proofs of Model (6)

The original method of Harvey and Shephard (1996) concerns the estimation of the following asymmetric SV model:

$$\begin{aligned} x_t &= \sigma \exp(h_t^*/2)\varepsilon_t \\ h_{t+1}^* &= \beta h_t^* + \eta_t \end{aligned} \qquad \begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim ID\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \gamma \\ \gamma & \sigma_\eta^2 \end{bmatrix} \right) \tag{A.1}$$

where ε_t is a process of independent, identically distributed random disturbances, symmetrically distributed around 0. The authors demonstrated that model (A.1) can be estimated applying the QML method to the time varying state space model:

$$\begin{aligned} y_t &= \omega^* + h_t^* + \xi_t \\ h_{t+1}^* &= \mu^* s_t + \beta h_t^* + \eta_t^* \end{aligned}$$
$$\begin{bmatrix} \xi_t \\ \eta_t^* \end{bmatrix} \mid s_t \sim ID\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_\xi^2 & \gamma^* s_t \\ \gamma^* s_t & \sigma_\eta^2 - \mu^{*2} \end{bmatrix} \right) \tag{A.2}$$

where:

- $y_t = \ln x_t^2$;
- $\omega^* = \ln \sigma^2 + E[\ln \varepsilon_t^2]$;
- $\xi_t = \ln \varepsilon_t^2 - E[\ln \varepsilon_t^2]$;
- $s_t = +1(-1)$ when x_t, i.e. ε_t, is positive (negative);
- μ^* and γ^* correspond to $E(\eta_t \mid s_t = 1)$ and $Cov(\eta_t, \varepsilon_t \mid s_t = 1)$ respectively.

Harvey and Shephard (1996) proved that the relations of μ^* and γ^* with γ depend on the distribution of (η_t, ε_t). When (η_t, ε_t) is bivariate normal, it results (approximately) that:

$$\mu^* = 0.80\gamma \quad \gamma^* = 1.11\gamma \quad \sigma_\xi^2 = 4.93 \tag{A.3}$$

Using the restrictions (A.3) and the transformation $h_t = h_t^* + 1.27 + \omega^*$, Model (A.2) assumes, after some algebra, the form (6):

$$\begin{aligned} y_t &= -1.27 + h_t + \xi_t \\ h_{t+1} &= \omega + 0.80\gamma s_t + \beta h_t + \eta_t^* \end{aligned}$$
$$\begin{bmatrix} \xi_t \\ \eta_t^* \end{bmatrix} \mid s_t \sim ID\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 4.93 & 1.11\gamma s_t \\ 1.11\gamma s_t & \sigma_\eta^2 - 0.64\gamma^2 \end{bmatrix} \right) \tag{A.4}$$

where $\omega = (1 - \beta)(1.27 + \omega^*)$.

Model (A.4) can also be represented in the alternative form:

$$\begin{aligned} y_t &= -1.27 + h_t + \mathbf{c}'\mathbf{u_t} \\ h_{t+1} &= \omega + 0.80\gamma s_t + \beta h_t + \mathbf{b_t}'\mathbf{u_t} \end{aligned} \qquad \mathbf{u_t} \mid s_t \sim ID(\mathbf{0}, \mathbf{I}) \tag{A.5}$$

where $\mathbf{c}' = [\sqrt{4.93}, \ 0]$, $\mathbf{b_t}' = [1.11\gamma s_t/\sqrt{4.93}, \ \sigma_{\eta^+}]$; σ_{η^+} is the standard deviation of the exogenous innovation on h_t.

Taking into account Formulas (3.2.3c), (3.2.4a) and (3.2.22) in the book of Harvey (1990), the one step ahead predictor of h_t is:

$$\begin{aligned}
\hat{h}_{t+1|t} &= \omega + \beta \hat{h}_{t|t-1} + 0.80\gamma s_t + \kappa_t(y_t - \hat{y}_{t|t-1}) \\
\kappa_t &= \frac{\beta P_{t|t-1} + 1.11\gamma s_t}{P_{t|t-1} + 4.93}
\end{aligned} \tag{A.6}$$

where: $\hat{y}_{t|t-1} = -1.27 + \hat{h}_{t|t-1}$ is the one step ahead prediction of y_t; $P_{t|t-1} = E[(\hat{h}_{t|t-1} - h_t)^2]$ is the MSE of $\hat{h}_{t|t-1}$.

Appendix B Proofs of Formulas (12) and (13)

A scaled Student's *t*-distribution, with v degrees of freedom, is the distribution of a random variable, ζ_v, defined as:

$$\zeta_v = \frac{Z}{\sqrt{\chi_v^2/v}}\sqrt{(v-2)/v} \tag{B1}$$

where Z is a standard normal r.v., and χ_v^2 is a chi-square r.v. with v degrees of freedom, independent of Z.

It is easy to derive that: $E(\zeta_v) = 0$ and $Var(\zeta_v) = 1$.

Given $\xi_v = \ln \zeta_v^2$, it follows:

$$\begin{aligned}
\xi_v &= \ln(\chi_1^2) - \ln(\chi_v^2/v) + \ln[(v-2)/v] \\
E[\xi_v] &= E[\ln(\chi_1^2)] - E[\ln(\chi_v^2/v)] + \ln[(v-2)/v] \\
Var[\xi_v] &= Var[\ln(\chi_1^2)] + Var[\ln(\chi_v^2/v)]
\end{aligned} \tag{B2}$$

From Zelen and Severo (1972) we know that:

$$\begin{aligned}
E[\ln(\chi_v^2/v)] &= \psi_0(v/2) - \ln(v/2) \\
Var[\ln(\chi_v^2/v)] &= \psi_1(v/2)
\end{aligned} \tag{B3}$$

where ψ_0 and ψ_1 are the *digamma* and *trigamma* function, respectively (Davis 1972).

Since $\psi_0(1/2) \approx -1.963$, $\psi_1(1/2) \approx 4.93$, and $\ln(1/2) \approx -0.693$, it follows approximately that:

$$\begin{aligned}
E[\xi_v] &= -1.963 + 0.693 - \psi_0(v/2) + \ln(v/2) + \ln[(v-2)/v] \\
&= -1.27 - \psi_0(v/2) + \ln(v/2 - 1) \\
Var[\xi_v] &= 4.93 + \psi_1(v/2)
\end{aligned} \tag{B4}$$

Data availibility The data used in the work can be requested directly from the author.

Declarations

Conflict of interest Author declare that there is no conflict of interest.

Consent to participate The work was carried out in compliance with the ethical code of the Università del Piemonte Orientale.

Code availability The gretl script for IQML can be requested directly from the author.

Consent for publication The author agrees for publication.

References

Black F (1976) Studies of stock market volatility changes. In: Proceedings of the American statistical association, business & economic statistics section

Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. J Econ 31:307–327

Christie AA (1982) The stochastich behaviour of common stock variances: value, leverage, and interest rate effects. J Finan Econ 10:407–432

Cottrell A, Lucchetti R (2021) Gretl user's guide. Distributed with the Gretl library

Davis P (1972) Chapter 6. Gamma function and related functions. In: Abramovitz M, Stegun IA (eds) Hanbook of mathematical functions. Dover, New York

Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc Ser B (Methodological) 39(1):1–22

Dunsmuir W (1979) A central limit theorem for parameter estimation in stationary vector time series and its application to models for a signal observed with noise. Ann Stat 490–506

Engle RF, Ng VK (1993) Measuring and testing the impact of news on volatility. J Financ 48(5):1749–1778

Geweke J (1986) Modelling the persistence of conditional variances: a comment. Econ Rev 5:57–61

Glosten L, Jagannathan R, Runkle D (1993) On the relation between the expected value and volatility of nominal excess returns on stocks. J Financ 46:1779–1801

Hamilton J (1994) Time series analysis. Princeton University Press, Princeton, NJ

Hansen PR, Lunde A (2005) A forecast comparison of volatility models: does anything beat a Garch(1,1)? J Appl Econ 20(7):873–889

Harvey A, Shephard N (1996) Estimation of an asymmetric stochastich volatility model for asset returns. J Bus Econ Stat 14(4):429–434

Harvey A, Ruiz E, Shepherd N (1994) Multivariate stochastic volatility methods. Rev Econ Stud 61:247–264

Harvey AC (1990) Forecasting, structural time series models and the Kalman filter. Cambridge University Press, Cambridge

Hull J, White A (1987) The pricing of options on assets with stochastic volatility. J Financ 42:281–300

Melino A, Turnbull SM (1990) Pricing foreign currency options with stochastic volatility. J Econom 45(1–2):239–265

Nelson D (1991) Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59:347–370

Pantula S (1986) Modelling the persistence of conditional variances: a comment. Econom Rev 5:71–73

Patton AJ (2011) Volatility forecast comparison using imperfect volatility proxies. J Econom 160(1):246–256

Ruiz E (1994) Quasi-maximum likelihood estimation of stochastic volatility models. J Econom 63(1):289–306

Sandmann G, Koopman SJ (1998) Estimation of stochastic volatility models via Monte Carlo maximum likelihood. J Econom 87(2):271–301

Shumway RH, Stoffer DS (1982) An approach to time series smoothing and forecasting using the EM algorithm. J Time Ser Anal 3(4):253–264

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

Taylor SJ (1994) Modeling stochastic volatility: a review and comparative study. Math Financ 4(2):183–204

Wiggins JB (1987) Option values under stochastic volatility: theory and empirical estimates. J Financ Econom 19(2):351–372

Zakoian JM (1994) Threshold heteroskedastic models. J Econom Dyn Control 18(5):931–955

Zelen M, Severo N (1972) Chapter 26. Probability function. In: Abramovitz M, Stegun IA (eds) Hanbook of mathematical functions. Dover, New York