

Finite sample properties of a QML estimator of stochastic volatility models with long memory

Ana Pérez^a, Esther Ruiz^{b,*}

^a*Department of Applied Economics (Statistics and Econometrics), Universidad de Valladolid, Valladolid, Spain*

^b*Department of Statistics and Econometrics, Universidad Carlos III de Madrid, Madrid, Spain*

Abstract

We analyse the finite sample properties of a QML estimator of LMSV models. We show up its poor performance for realistic parameter values. We discuss an identification problem when the volatility has a unit root. An empirical analysis illustrates our findings. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Fractional integration; Stochastic volatility; Quasi-maximum likelihood estimator

1. Introduction

High frequency financial time series of returns are often characterised by having excess kurtosis and autocorrelated squared observations. Moreover, these autocorrelations tend to decay very slowly, suggesting that squared returns could be characterised as a long memory process; see Ding et al. (1993). This property was incorporated into Stochastic Volatility (SV) models by Harvey (1998); Breidt et al. (1998), who propose Long-Memory SV (LMSV) models and suggest estimating the parameters of these models by Quasi-Maximum Likelihood (QML) in the frequency domain. Breidt et al. (1998) prove the strong consistency of this estimator and analyse its finite sample properties for some of the parameters of the model in the stationary case concluding that the estimator behaves properly. However, the parameter values they consider in their simulations are far from the parameter values often estimated with real financial series. In this paper, we extend their study to all the parameters of the model and to a bigger range of parameter values that include more realistic and non-stationary cases and conclude that the properties of the QML estimator may be poor in some

*Corresponding author.: Tel.: +34-916-249-851; fax: +34-916-249-849.

E-mail address: ortega@est-econ.uc3m.es (E. Ruiz).

cases of interest. Finally, we illustrate the problems faced when estimating LMSV models analysing a series of daily returns for the IBEX35 index.

2. QML estimation of LMSV models

The AutoRegressive Long Memory Stochastic Volatility (ARLMSV) model is given by:

$$y_t = \sigma_* \sigma_t \varepsilon_t \quad (1a)$$

$$(1 - \phi L)(1 - L)^d \log(\sigma_t^2) = \eta_t \quad (1b)$$

where σ_t denotes the volatility of series y_t , σ_* is a scale parameter, ε_t is a Gaussian white noise process with unit variance, η_t is NID(0, σ_η^2) and independent of ε_t , and d can be non-integer.

Model (1a,b) encompasses several SV models previously proposed in the literature. When $\phi=0$, model (1a,b) becomes the basic LMSV model in Harvey (1998); if $d=0$ and $|\phi|<1$, we obtain the ARSV (AutoRegressive SV) model in Harvey et al. (1994); when $\{d=0, \phi=1\}$ or $\{\phi=0, d=1\}$, model (1a,b) becomes the RWSV (Random Walk SV) model. Notice that in the last case, the parameters of the ARLMSV model are not identified.

Model (1a,b) can be linearised as follows:

$$\log(y_t^2) = \mu + \log(\sigma_t^2) + \xi_t, \quad (2)$$

where $\mu = \log(\sigma_*^2) + E(\log(\varepsilon_t^2))$ and $\xi_t = \log(\varepsilon_t^2) - E(\log(\varepsilon_t^2))$. The QML estimator proposed by Harvey (1998) and Breidt et al. (1998) is obtained by maximising the discrete Whittle approximation to the likelihood function of $\log(y_t^2)$ in the frequency domain, treating ξ_t as if it were Gaussian. When $-0.5 < d < 0.5$ and $|\phi| < 1$, the model is stationary and the QML estimator is consistent; see Breidt et al. (1998). The asymptotic distribution of QML estimators for LMSV models is still unknown. If $0.5 \leq d \leq 1$, the model is not stationary, but taking first differences in Eq. (2) yields a stationary series and QML estimation can then be carried out for $\Delta \log(y_t^2)$.

Once the parameters of the model have been estimated, the underlying volatility at time t may be estimated by $\hat{\sigma}_* \exp\{0.5 \tilde{h}_{t/T}\}$, where $\hat{\sigma}_*$ is the sample standard deviation of the heteroscedasticity corrected observations, as suggested by Harvey and Shephard (1993), and $\tilde{h}_{t/T}$ is the smoothed estimated log-volatility series obtained using the algorithm in Harvey (1998).

3. Finite sample properties of QML estimator

In this section, we analyse the finite sample properties of the Whittle estimator of the parameters of the ARLMSV model by means of several Monte Carlo experiments. As ε_t is assumed to be Gaussian, the variance of ξ_t is known to be $\pi^2/2$ and is not estimated. For each model design and sample size considered, $T=1024$, $T=4096$ and $T=8192$, we generate 3000 replicates. Table 1 displays the bias and standard deviations for some selected ARLMSV models.

In the basic stationary LMSV models ($\phi=0, d<0.5$), we observe that the QML estimator of d has a large negative bias and its variance decreases as d increases. On the other hand, the estimator of σ_η^2 is

Table 1

Simulation bias and standard deviation (in parenthesis) for the spectral QML estimators under several ARLMSV models

Parameters (ϕ, d, σ_η^2)	ϕ			d			σ_η^2		
	$T=1024$	$T=4096$	$T=8192$	$T=1024$	$T=4096$	$T=8192$	$T=1024$	$T=4096$	$T=8192$
(0, 0.2, 0.1)	–	–	–	–0.026 (0.171)	–0.027 (0.157)	0.000 (0.148)	0.121 (0.265)	0.056 (0.149)	0.036 (0.115)
(0, 0.4, 0.5)	–	–	–	–0.009 (0.094)	–0.002 (0.052)	–0.001 (0.037)	0.069 (0.330)	0.030 (0.171)	0.021 (0.124)
(0, 0.4, 0.1)	–	–	–	–0.100 (0.178)	–0.026 (0.111)	–0.011 (0.081)	0.137 (0.265)	0.061 (0.137)	0.035 (0.095)
(0, 0.45, 0.1)	–	–	–	–0.101 (0.161)	–0.032 (0.086)	–0.017 (0.063)	0.141 (0.258)	0.060 (0.124)	0.034 (0.083)
(0, 0.49, 0.1)	–	–	–	–0.105 (0.141)	–0.043 (0.068)	–0.027 (0.048)	0.149 (0.250)	0.064 (0.111)	0.039 (0.071)
(0, 0.75, 0.1)	–	–	–	0.044 (0.152)	0.020 (0.081)	0.015 (0.056)	0.053 (0.156)	0.012 (0.064)	0.004 (0.040)
(0.2, 0.4, 0.1)	0.167 (0.520)	0.073 (0.478)	0.027 (0.425)	–0.136 (0.189)	–0.070 (0.131)	–0.034 (0.092)	0.116 (0.271)	0.057 (0.156)	0.041 (0.121)
(0.2, 0.45, 0.1)	0.174 (0.530)	0.072 (0.471)	0.026 (0.424)	–0.161 (0.181)	–0.065 (0.117)	–0.032 (0.081)	0.120 (0.271)	0.059 (0.155)	0.044 (0.123)
(0.9, 0.2, 0.1)	–0.052 (0.190)	–0.013 (0.089)	–0.003 (0.050)	–0.024 (0.171)	–0.020 (0.122)	–0.016 (0.088)	0.071 (0.177)	0.024 (0.071)	0.013 (0.041)
(0.99, 0.2, 0.1)	–0.022 (0.019)	–0.018 (0.010)	–0.018 (0.007)	0.224 (0.139)	0.275 (0.070)	0.286 (0.049)	–0.065 (0.052)	–0.087 (0.018)	–0.090 (0.013)

positively biased. Regarding the non-stationary model with $d=0.75$, it seems that the QML estimator behaves better than in the stationary cases, with small standard deviations and smaller, but positive, bias for \hat{d} and smaller bias for $\hat{\sigma}_\eta^2$. Fig. 1 represents kernel estimates of the simulated probability densities of the QML estimates of d and σ_η^2 for models with $\{d=0.4, \sigma_\eta^2=0.1\}$, $\{d=0.4, \sigma_\eta^2=0.5\}$ and $\{d=0.75, \sigma_\eta^2=0.1\}$. For the first two models, the QML estimators are not adequate at all when $T=1024$. The distribution of \hat{d} has always a concentration of mass at $d=0.5$ and the situation becomes not much better, even with $T=8192$, when $\sigma_\eta^2=0.1$. For this model, we also observe that the sample distribution of $\hat{\sigma}_\eta^2$ is shifted towards zero, so that we will estimate $\hat{\sigma}_\eta^2 \approx 0$ with high probability. The estimation only improves considerably with the sample size when $\sigma_\eta^2=0.5$, that is a value clearly bigger than the values often estimated with real data. Regarding the LMSV model with $d=0.75$, the empirical distributions are such that inference is reliable, even when $T=4096$.

In the ARLMSV models ($0 < d < 0.5$, $|\phi| < 1$), there is always a negative bias in \hat{d} and a positive bias in $\hat{\sigma}_\eta^2$, both decreasing the bigger is T , except for the model with $\phi=0.99$. In this case, the bias of \hat{d} is positive and quite large and the bias of $\hat{\sigma}_\eta^2$ is negative, and none of them decreases with T . When the parameter d is near the boundary of non-stationarity ($d=0.5$), the parameter ϕ is largely overestimated and has much dispersion, while d is always underestimated. It seems that the QML estimator cannot distinguish whether the non-stationarity comes from the autoregressive polynomial ($\phi=1$) or from the fractional integration ($d=0.5$). This could be expected as the parameters of the ARLMSV model are not identified when the volatility process has a unit root. Fig. 2 displays the results for two ARLMSV models, $\{\phi=0.99, d=0.2, \sigma_\eta^2=0.1\}$ and $\{\phi=0.2, d=0.4, \sigma_\eta^2=0.1\}$. When

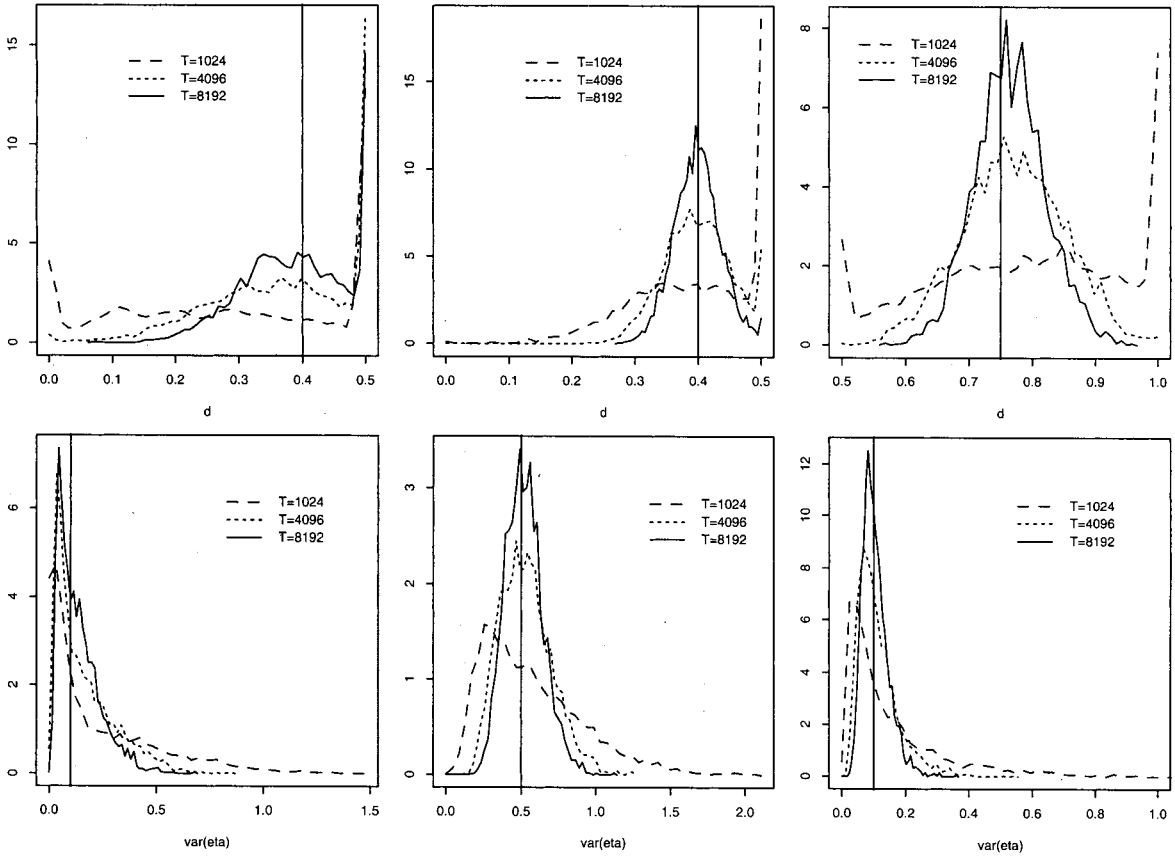


Fig. 1. Estimated densities of spectral QML estimators of d (top panels), and σ_η^2 (bottom panels) for three LMSV models, left to right: $\{d=0.4, \sigma_\eta^2=0.1\}$, $\{d=0.4, \sigma_\eta^2=0.5\}$, $\{d=0.75, \sigma_\eta^2=0.1\}$. Vertical line drawn on each graph indicates the true parameter value.

$\phi=0.99$, none of the parameters of the ARLMSV model can be estimated with enough precision. When $d=0.4$, the distribution of $\hat{\phi}$ has a concentration of mass close to one greater than expected, explaining the positive and large bias reported in Table 1.

Therefore, the finite sample properties of the QML estimator are such that inference on the parameters of ARLMSV models based on it may not be reliable. This conclusion contradicts the results in Breidt et al. (1998). However, they only consider $\sigma_\eta^2=1$ and $d=\{0.2, 0.4\}$, where the QML estimator behaves quite well.

4. Empirical analysis of IBEX35

Fig. 3 represents daily returns of the IBEX35 index of the Madrid Stock Exchange from 7/1/87 to 30/12/98 (2991 observations) after having removed some small correlation structure in the mean. The kurtosis of this series is 8.321 and the Box–Ljung statistic for the first ten autocorrelations of the squares is $Q_2(10)=1129.1$, which is highly significant. The squared and log-squared returns have

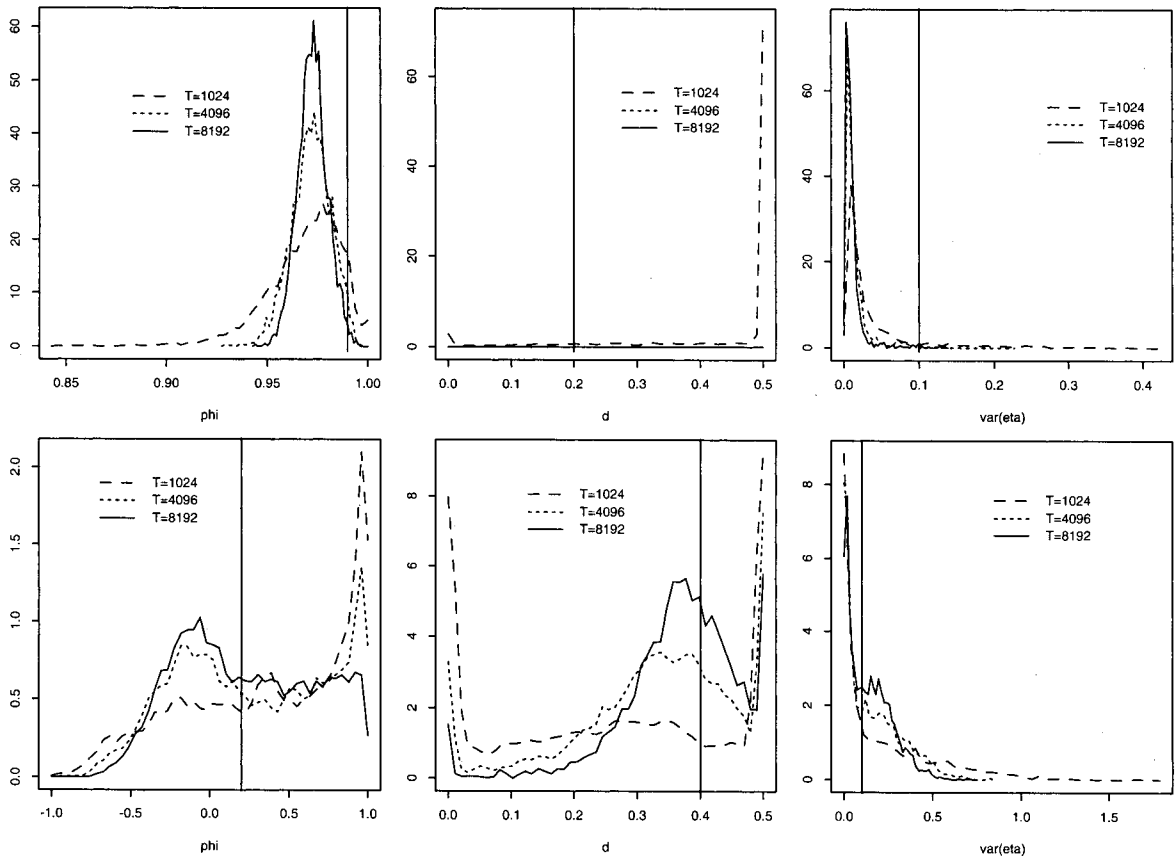


Fig. 2. Estimated densities of spectral QML estimators of ϕ (left), d (center) and σ_η^2 (right) for ARLMSV models with parameters $\{\phi=0.99, d=0.2, \sigma_\eta^2=0.1\}$ (top panels) and $\{\phi=0.2, d=0.4, \sigma_\eta^2=0.1\}$ (bottom panels). Vertical line drawn on each graph indicates the true parameter value.

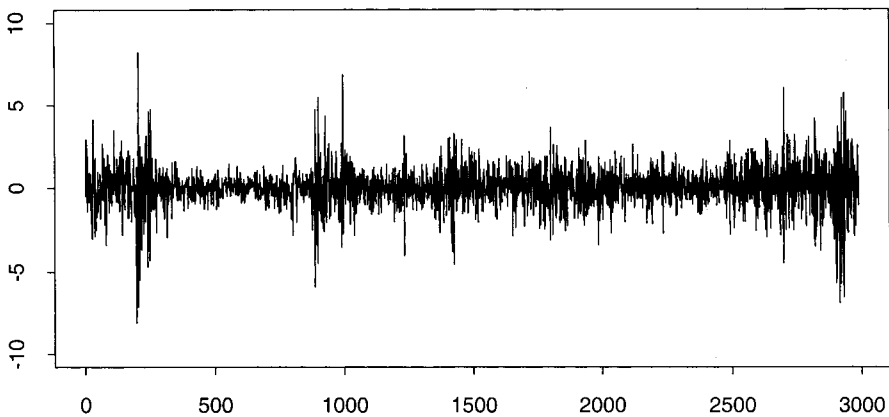


Fig. 3. Returns of IBEX35 observed daily from 7/1/87 to 30/12/98.

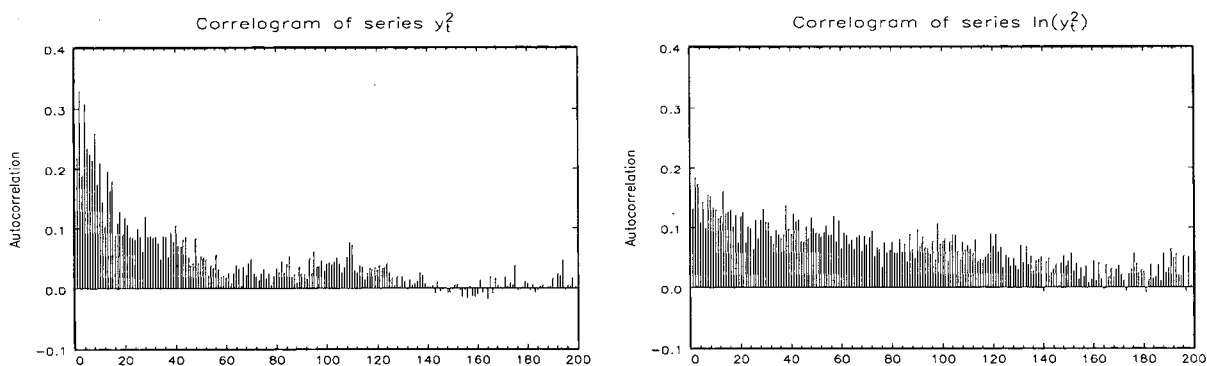


Fig. 4. Sample autocorrelations of squared returns (left) and log-squared returns (right) of IBEX35.

significant autocorrelations, even for very high lags, with a very slow decay, especially remarkable in the log-squared returns; see Fig. 4.

Consequently, we have first estimated the ARSV model by the QML method of Harvey et al. (1994). The estimated parameters together with their standard deviations, computed using the formulae in Ruiz (1994) and Harvey and Shephard (1993), appear in Table 2. As expected, the autoregressive parameter is estimated very close to one. The results of fitting a RWSV model are also reported in Table 2.

Regarding long memory models, we have first fitted a LMSV model, obtaining an estimate of d on

Table 2

Estimated parameters (standard deviation in parenthesis) and sample moments of observations standardised using smoothed volatility from several ARLMSV models^{a,b}

Model	ARSV	RWSV	LMSV	ARLMSV
Estimated Parameters	$\hat{\sigma}_\eta^2 = 0.0168$ (0.0057)	$\hat{\sigma}_\eta^2 = 0.0099$ (0.0024)	$\hat{\sigma}_\eta^2 = 0.0906$	$\hat{\sigma}_\eta^2 = 0.0155$
	$\hat{\phi} = 0.9898$ (0.0042)	-	-	$\hat{\phi} = 0.6632$
	$\hat{\sigma}_* = 0.9297$ (0.0374)	$\hat{\sigma}_* = 0.9484$ (0.0436)	$\hat{d} = 0.7538$ $\hat{\sigma}_* = 1.5112$	$\hat{d} = 0.7035$ $\hat{\sigma}_* = 1.5484$
Mean	0.0101	0.0074	0.0113	0.0122
Variance	1.0000	1.0000	0.9999	0.9999
Skewness	-0.0284	-0.0441	-0.0183	-0.0127
Kurtosis	3.5620	3.7070	3.3941	3.4380
$Q_2(10)$	43.15	80.06	17.89	15.50
$Q_2(50)$	101.70	144.61	71.37	68.43
$Q_2(100)$	150.41	198.87	117.67	113.39

^a $Q_2(k)$ denotes the Ljung-Box statistic for the first k autocorrelations in the squared observations.

^b No standard deviation of QML estimators is displayed for the long memory models, because their asymptotic distribution is still unknown.

the boundary of non-stationarity ($\hat{d}=0.5$). As Harvey (1998) suggests this could be indicating that a non-stationary model is preferable, so a model with $d>0.5$ has been estimated. Finally, we have also fitted an ARLMSV model. The results appear in Table 2.

Several sample moments of the standardised observations from the models previously estimated are also displayed in Table 2. The fit from long memory models is better than from the other models, with correlations in the squares being no longer significant at the 1% level. However, looking at the diagnostics on the residuals, it is difficult to distinguish between both long memory models. Notice that the implications of both models in terms of forecasting future volatilities are different because the variance of the unpredictable component in the LMSV model is bigger than in the ARLMSV model, which has an AR component that could be predicted.

In order to choose between both long memory models, and given that the asymptotic distribution of the QML estimators for these models is still unknown, we have carried out a likelihood ratio test to gauge the significance of the AR component. Recall that the ARLMSV model nests the basic LMSV model ($\phi=0$). The value of the test statistic is $LR = 2[\log L(\hat{\psi}) - \log L(\hat{\psi}_0)] = 1.0885$, which does clearly not reject the null compared to the χ_1^2 distribution. Therefore, the LMSV model seems to be preferable.

In Fig. 5, we plot the implied smoothed volatility of the IBEX35 returns for the four estimated models. As expected there is no difference between the volatilities estimated by the ARSV and the RWSV models. Moreover, the volatility implied by the two long memory models also show the same pattern. Comparing the long memory volatilities with the volatility implied by the ARSV and the

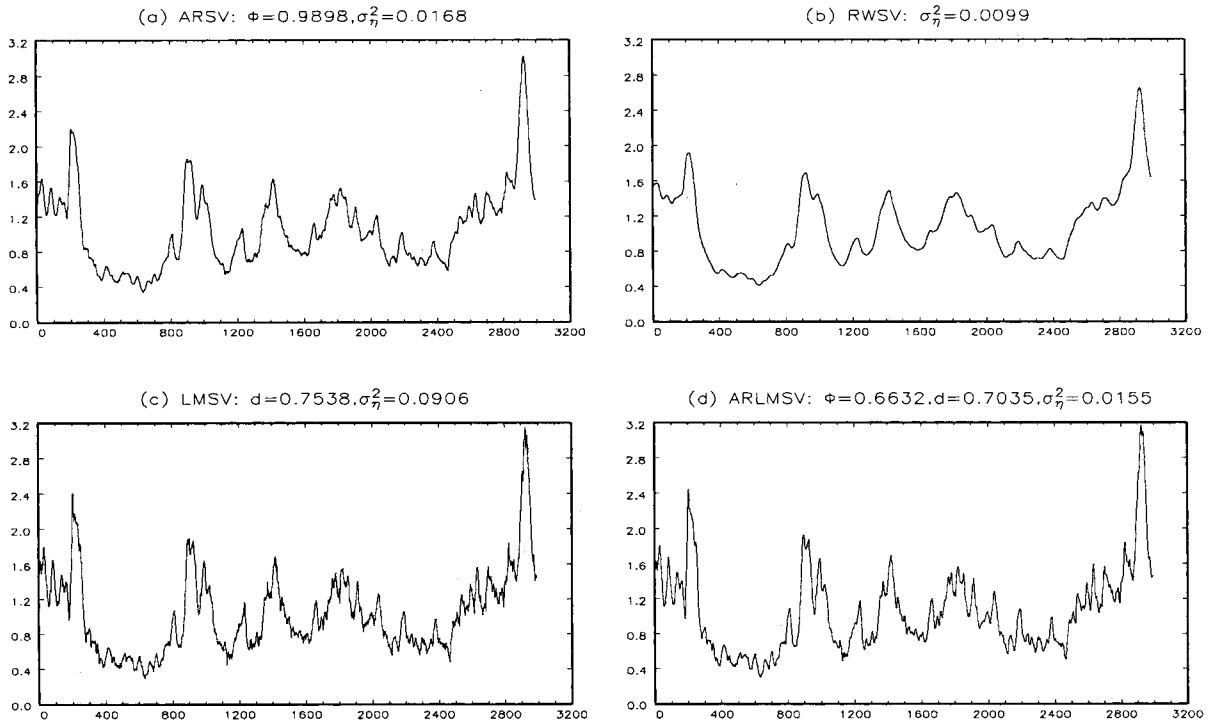


Fig. 5. Smoothed estimates of IBEX35 volatility from several SV models.

RWSV models, it is possible to observe that the fit of the former is better, especially in the high volatility periods.

5. Conclusions

We have shown that the finite sample properties of the spectral QML estimator for ARLMSV models are poor when the volatility evolves smoothly over time and/or is close to be non-stationary. Even worse, in the unit root case, the parameters of the ARLMSV model are not identified so inference is not reliable at all. When modelling the volatility of the IBEX35 index, it seems that this can be represented by a long memory process but it is difficult to establish the adequate specification of that process.

Acknowledgements

This research was started while Ana Pérez was a visitor at the Department of Statistics and Econometrics (Universidad Carlos III de Madrid). She thanks the department for its hospitality. We are very grateful to A.C. Harvey, J. Hidalgo, F. Marmol, F. Sowell, C. Velasco, and the participants at the 1999 Time Series Workshop held in Arrabida, Portugal, specially to B. Ray and N. Crato, for helpful comments and suggestions. We also acknowledge financial support from projects SEC97-1379 and PB95-0299 by the Spanish Government. Any remaining errors are our own.

References

- Breidt, F.J., Crato, N., de Lima, P.J.F., 1998. The detection and estimation of long-memory in stochastic volatility. *Journal of Econometrics* 83, 325–348.
- Ding, Z., Granger, C.W.J., Engle, R.F., 1993. A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1, 83–106.
- Harvey, A.C., 1998. Long memory in stochastic volatility. In: Knight, J., Satchell, E. (Eds.), *Forecasting Volatility in Financial Markets*. Butterworth–Haineman, London, pp. 307–320.
- Harvey, A.C., Ruiz, E., Shephard, N.G., 1994. Multivariate stochastic variance models. *Review of Economic Studies* 61, 247–264.
- Harvey, A.C., Shephard, N.G., 1993. Estimation and testing of stochastic volatility models. *STICERD Econometrics Discussion Paper EM/93/268*, London School of Economics.
- Ruiz, E., 1994. Quasi-maximum likelihood estimation of stochastic volatility models. *Journal of Econometrics* 63, 289–306.