A Computationally Efficient Method for Obtaining Smoothed Volatilities in Long-Memory Stochastic Volatility Models

Francesc Marmol[1], **Ana Pérez**[2], **Juan C. Reboredo**[3]

[1] *Department of Economics, Universidad Carlos III de Madrid, Spain*

[2] *Department of Applied Economics, Universidad de Valladolid, Spain*

[3] *Department of Economics, Universidade de Santiago de Compostela, Spain*

Abstract We provide a computationally efficient method, based on Harvey (1998) proposal, to estimate the underlying volatility of asset returns using the Long-Memory Stochastic Volatility ($LMSV$) model. The performance of our procedure is illustrated with an application to three series of daily exhange rates returns. A comparison of long memory GARCH-type volatilities with our smoothed ones is also presented.

Keywords Stochastic Volatility, Long-Range Dependence, Smoothing, Exchange Rate, FIEGARCH.

JEL Classification C22, C52, G10.

1. Introduction

It is a well established fact that many financial series such as asset returns are serially uncorrelated over time, but display a highly persistent behavior in some

Correspondence to: Ana Pérez (e-mail: `perezesp@eaee.uva.es`)

nonlinear transformations such as squared, log-squared and absolute returns. Indeed, a large body of empirical research has documented that the autocorrelations of the return volatility series decay at a slow hyperbolic rate instead of exponentially like in weakly stationary models; see *e.g.* Dacorogna *et al.* (1993), Ding *et al.* (1993), Bollerslev and Mikkelsen (1996), Bollerslev and Wright (2000) and Andersen *et al.* (2001a, 2001b), among others. This empirical fact motivated Breidt *et al*. (1998) and Harvey (1998) to introduce the Long Memory Stochastic Volatility (*LMSV*) model.

The $LMSV$ model for a time series of returns y_t can be defined as

$$y_t = \sigma_t \varepsilon_t,\ \ \sigma_t = \sigma \exp(h_t/2), \tag{1}$$

where σ_t is the volatility, $\varepsilon_t \sim IID(0,1)$, $\sigma > 0$ and h_t is a Gaussian $ARFIMA(p, d, q)$ process independent of ε_t given by

$$\phi(B)(1-B)^d h_t = \theta(B)\eta_t, \tag{2}$$

where B is the lag operator, $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ are stationary autoregressive and invertible moving average operators, respectively, d is the fractional differencing parameter and $\eta_t \sim NID(0, \sigma_\eta^2)$. If $d < 1/2$, the process y_t is both covariance and strictly stationary and it is nonstationary otherwise. However, if $1/2 \leq d < 1$, the volatility process has transitory memory, *i.e.*, any random shock has only temporary influence on the volatility series, whereas if $d \geq 1$, the shocks persist indefinitely on its future path.

An appealing feature of the *LMSV* model is that volatility does not only depend on past observations as in *ARCH* models, but also on some stochastic unobserved components. However, its estimation is not an easy task, because the *LMSV* model is not conditionally Gaussian and, therefore, the exact likelihood is very difficult to evaluate. Nevertheless, several estimation procedures are now available in the econometric literature. These include parametric methods, such us the Frequency Domain Quasi Maximum Likelihood (*FDQML*) method, independently proposed by Breidt *et al.* (1998) and Harvey (1998), the enhanced *FDQML* method recently proposed by Deo *et al*. (2006), the Generalised Method of Moments (*GMM*) suggested by Wright (1999a) and the Bayesian approach of So (1999, 2002). The *LMSV* model's fractional differencing parameter, d, can also be estimated with semiparametric methods; see, for example, Deo and Hurvich (2001), Hurvich and Ray (2003), Arteche (2004), Jensen (2004) and Hurvich *et al*. (2005). However, the semiparametric methods do not provide estimators for the other parameters of the *LMSV* model, but only for d. This rules them out from being used as the basis for the smoothing algorithm to be explained in the following sections.

Once a *LMSV* model has been estimated, the natural course of action is to obtain estimates of the underlying volatility as it has important and practical implications for asset pricing, portfolio allocation, and risk management. From a statistical point of view, the estimation of the volatility process is essential for diagnostic checking and model selection based on standardized residuals. Harvey (1998) proposes a method to carry out exact smoothing and to obtain an estimator of the long-memory signal that requires the inversion of a $n \times n$ matrix, n denoting the available sample size. Since the inversion of this matrix could become cumbersome for large sample sizes, Harvey suggests to do the smoothing using weights worked out for a smaller sample size and argues that little accuracy would be lost because the weights given to the remote observations are very small. In this paper we provide a feasible implementation of Harvey's proposal, and obtain an estimation of the long memory signal that renders accurate volatility estimations[1].

[1] Chen *et al*. (2006) propose a method for inverting Toeplitz matrices using a version of the preconditioned conjugate gradient method that can be applied to LMSV models to obtain the exact volatilities for very large data sets. How this proposal compares to ours is a topic left for further research.

The rest of the paper proceeds as follows. The next section becomes precise on the Harvey's smoothing algorithm and introduces our feasible version. Section 3 presents an application to the modelling of daily exchange returns that includes a comparison of our smoothed volatilities with those estimated by long-memory GARCH-type models. Finally, Section 4 concludes.

2. Signal Extraction and Smoothing

One of the main goals in fitting a model for time varying volatilities of a financial time series is to obtain an estimation of the volatility itself. In the *LMSV* model, Harvey (1998) proposed a method to carry out exact smoothing and obtain an estimator of the signal, h_t, based on the full sample. We summarize here the most relevant aspects of this method and then we present our feasible solution when the sample size, n, is too large to enable Harvey's original proposal to be implemented.

The smoothing procedure proposed by Harvey (1998) exploits the fact that the *LMSV* model defined in (1)-(2) can be written in an equivalent way for the log-squared returns, as:

$$x_t = \ln(y_t^2) = \mu + h_t + \xi_t, \tag{3}$$

where $\mu = \ln(\sigma^2) + E(\ln(\varepsilon_t^2))$ and $\xi_t = \ln(\varepsilon_t^2) - E(\ln(\varepsilon_t^2))$. From (2)-(3), it turns out that the *LMSV* model can be thought of as a structural time series model where x_t is given by the sum of a Gaussian *ARFIMA* signal h_t plus an *i.i.d.* non-Gaussian noise ξ_t with zero mean and variance σ_ξ^2.

Equation (3) can be written in matrix notation as follows:

$$\mathbf{x} = \mu\,\mathbf{i} + \mathbf{h} + \xi,$$

where $\mathbf{x}$ is a $n \times 1$ vector containing the observations x_t, for $t = 1, 2, ..., n$; μ is the mean of the process x_t; $\mathbf{i}$ is a $n \times 1$ vector of ones; $\mathbf{h}$ is a $n \times 1$ vector containing the signal values; and ξ is a $n \times 1$ vector with the disturbance values. When h_t is a stationary process, the *Minimum Mean Square Linear Estimator* (*MMSLE*) of h_t that provides the smoothed estimation of $\mathbf{h}$ is given by the following expression

$$\widetilde{\mathbf{h}} = (\mathbf{I} - \sigma_\xi^2 \mathbf{V}^{-1})\mathbf{x} + \sigma_\xi^2 \mathbf{V}^{-1}\mathbf{i}\mu = \mathbf{x} - \sigma_\xi^2 \mathbf{V}^{-1}(\mathbf{x} - \mu\,\mathbf{i}), \tag{4}$$

where μ is an unknown parameter that will be estimated by the sample mean of x_t, $\mathbf{V}$ denotes the covariance matrix of x_t and $\mathbf{I}$ is the $n \times n$ identity matrix. Since the process ξ_t is white noise $(0, \sigma_\xi^2)$ and independent of h_t, it turns out that $\mathbf{V} = \mathbf{V}_h + \sigma_\xi^2 \mathbf{I}$, where $\mathbf{V}_h$ is the covariance matrix of h_t. In the particular case where h_t is a stationary $ARFIMA(0, d, 0)$ process, the ij-th element of $\mathbf{V}_h$ is $\gamma_h(k) = \{(k-1+d)/(k-d)\}\,\gamma_h(k-1)$, with $k = |i-j|$, $k = 1, 2, \ldots$, and $\gamma_h(0) = \sigma_h^2 = \sigma_\eta^2 \mathit{\Gamma}(1-2d)/\left\{\mathit{\Gamma}(1-d)\right\}^2$ (see Hosking (1981)). In the general stationary $ARFIMA(p, d, q)$ case, the autocovariances are complicated functions of hypergeometric functions but they can be efficiently computed with the algorithm in Chung (1994) and the more recent one proposed by Bertelli and Caporin (2002).

When the process h_t is nonstationary ($1/2 \leq d < 1$) the smoothing formulae applies to the model in first differeces, given by:

$$\triangle x_t = \triangle h_t + \triangle \xi_t,$$

where $\triangle h_t$ is a stationary $ARFIMA(p, d^*, q)$, with $d^* = d-1$ and $-1/2 \leq d^* < 0$. In this case, the term involving the mean μ vanishes and the algorithm would provide smoothed estimators of $\triangle h_t$, with $t = 2, \ldots n$. If we denote by $\widetilde{\mathbf{h}}^*$ the $(n-1) \times 1$ vector containing such estimators, the smoothing equation for $\triangle h_t$ can be written as

$$\widetilde{\mathbf{h}}^* = (\mathbf{I} - \mathbf{V}_\xi^* \mathbf{V}^{*-1})\mathbf{x}^*, \tag{5}$$

where $\mathbf{x}^*$ is a vector with the first differenced observations $\triangle x_t$, and $\mathbf{V}^*$ and $\mathbf{V}^*_\xi$ denote the covariance matrices of $\triangle x_t$ and $\triangle \xi_t$, respectively. Smoothed estimators of the signal h_t are then computed from the recursion:

$$\widetilde{h}_t = \widetilde{h}_{t-1} + \widetilde{h}^*_t, \; t = 2, \ldots n, \qquad \widetilde{h}_1 = 0.$$

Finally, the smoothed estimator of the underlying volatility, denoted by $\hat{\sigma}_t$, is obtained as the product:

$$\hat{\sigma}_t = \hat{\sigma} \; \exp(\widetilde{h}_t/2), \; t = 1, \ldots n$$

where $\hat{\sigma}$ is the estimator of the scale factor suggested by Harvey and Shephard (1993), i.e., $\hat{\sigma} = (n^{-1} \sum_{t=1}^{n} \widetilde{y}_t)^{1/2}$, with $\widetilde{y}_t = y_t \exp(-\widetilde{h}_t/2)$.

Note that this smoothing procedure requires the inversion of the $n \times n$ matrix $\mathbf{V}$ in equation (4) and this could pose computational problems when dealing with large data sets. To overcome this drawback, we propose to build up an approximate weight $n \times n$ matrix, say $\mathbf{W}$, alternative to the matrix $\mathbf{V}^{-1}$, that only requires computing the weights for a smaller sample size, say N ($N < n$). Our procedure is as follows[2]. First, let us define $\mathbf{V}_{h,N}$ as the $N \times N$ covariance matrix of $(h_1, \ldots, h_N)$. Since this matrix is symmetric Toeplitz, only N different elements, $\gamma_h(0), ..., \gamma_h(N-1)$, are needed to complete it and these can be evaluated in $O(N \log N)$ operations with the algorithm in Bertelli and Caporin (2002). We next obtain the covariance matrix of $(x_1, \ldots, x_N)$ as $\mathbf{V}_N = \mathbf{V}_{h,N} + \sigma^2_\xi \mathbf{I}$ and compute its inverse, i.e. $\mathbf{V}_N^{-1}$. Let ω_{ij} be the ij-th element of this matrix, with $i = 1, ..., N$, $j = 1, ..., N$, where $\omega_{ij} = \omega_{ji}$. In order to build up the new $n \times n$ matrix $\mathbf{W}$, we propose to shift this $N \times N$ matrix $\mathbf{V}_N^{-1}$ along the main diagonal of $\mathbf{W}$, and complete the off-diagonal elements of this matrix up to n with zeros. Then, the approximated smoothed values of the log-volatility, say $\widehat{\mathbf{h}}$, will

[2] We only explain our proposal for the stationary case. In the nonstationary case, a similar procedure is applied to the matrix $\mathbf{V}^*$ in equation (5).

be calculated as:

$$\widehat{\mathbf{h}} = (\mathbf{I} - \sigma_\xi^2 \mathbf{W})\mathbf{x} + \sigma_\xi^2 \mathbf{W}\mathbf{i}\overline{x} = \mathbf{x} - \sigma_\xi^2 \mathbf{W}(\mathbf{x} - \overline{x}\mathbf{i})$$

where $\overline{x}$ is the sample mean of $(x_1, ..., x_n)$. Note that this scheme only requires the inversion of the $N \times N$ matrix $\mathbf{V}_N$ and, since this is a symmetric Toeplitz matrix, the number of operations required to invert it is $O(N \log^2 N)$.

To better illustrate our procedure, consider a simple example where $n/3$ is an integer number and $n-N$ is even. In this case, the $n \times n$ matrix $\mathbf{W}$ could be build up by shifting the $N \times N$ matrix $\mathbf{V}_N^{-1}$ in three blocks along the main diagonal and completing the off-diagonal elements up to n with zeros. In the zones where the blocks overlap, we choose the elements from the closest block to the main diagonal in order to assign nonzero weights to the closest N observations in time. In particular, to get the approximated smoothed value of the volatility at time i, say $\widehat{h}_i$, with $i = 1, ..., n/3$, the weights assigned to the first N observations are the N elements of the i-th row of $\mathbf{V}_N^{-1}$ while the left observations $(x_{N+1}, \ldots, x_n)$' get zero weight. That is, the i-th row of the $n \times n$ matrix $\mathbf{W}$, for $i = 1, 2, ..., n/3$, will be:

$$\left(\underbrace{\omega_{i1}, ..., \omega_{ii}, ..., \omega_{iN}}_{N}, \underbrace{0, ..., 0}_{(n-N)} \right).$$

When $i = n/3 + 1, ..., 2n/3$, the first and last observations $(x_1, \ldots, x_{(n-N)/2})$ and $(x_{(n+N)/2+1}, \ldots, x_n)$ are given zero weight and the N central observations are multiplied with the corresponding central row of $\mathbf{V}_N^{-1}$. Therefore, for $i = n/3 + 1, ..., 2n/3$, the i-th row of $\mathbf{W}$ will be:

$$\left(\underbrace{0, ..., 0}_{(n-N)/2}, \underbrace{\omega_{i-(n-N)/2,1}, ..., \omega_{i-(n-N)/2,i-(n-N)/2}, ..., \omega_{i-(n-N)/2,N}}_{N}, \underbrace{0, ..., 0}_{(n-N)/2} \right).$$

Finally, for $i = 2n/3 + 1, ..., n$, the first $n - N$ observations get zero weight and the last N observations $(x_{n-N+1}, \ldots, x_n)$ are weighted with the elements of the corresponding bottom row of $\mathbf{V}_N^{-1}$. That is, for $i = 2n/3+1, ..., n$, the i-th row of $\mathbf{W}$ will be:

$$\left(\underbrace{0, ..., 0}_{(n-N)}, \underbrace{\omega_{N+i-n,1}, ..., \omega_{N+i-n,N+i-n}, ..., \omega_{N+i-n,N}}_{N} \right).$$

At this point there are two issues that deserve further discussion. First, the selection of the size N, and second, how to shift the NxN submatrix $\mathbf{V}_N^{-1}$ across the n observations. As regards the sample size N, it is clear that the larger we select N, the better the approximation, but the larger the number of multiplications required. In this sense, experiments in Pérez (2002) reveal that very reasonably results are obtained when choosing $N/n \in (0.5, 0.7)$. But even smaller values of N could lead to satisfactory results, as far as the remote elements in the rows of $\mathbf{V}^{-1}$ are very close to zero. Therefore, the choice of the truncation parameter N will depend mainly on the estimated parameter values: for very persistent estimated models with d close to $1/2$, N should be larger than for less persistent models where the extreme elements of the inverse matrix $\mathbf{V}^{-1}$ are very close to zero.

Another important concern is how to determine the number of overlapping block matrices needed to best construct the $n \times n$ matrix $\mathbf{W}$. Clearly, there are many possibilities, from shifting the matrix in few blocks overlapping as little as possible, to the other extreme of a more dense overlap that shifts the matrix very frequently. Which one constitutes an optimal approach is not investigated in this paper, but we just try overlappings that have shown to provide reasonable results in our simulation experiments.

As an illustration of the accuracy of our procedure, Figure 1 compares the implied weights for $\widetilde{h}_i$ in a very persistent $LMSV$ model with $\{d = 0.45$, $\sigma_\eta^2 = 0.1$, $\sigma = 1\}$ and $n = 840$ (solid line) with the approximate weights for $\widehat{h}_i$ with $N = 560$ (dots and dashes), $N = 420$ (closely dots) and $N = 350$ (short dashes). The left hand side panel displays the weights at time $i = 140$, while the right hand side panel corresponds to $i = 400$. For $N = 350$ and $N = 420$, the $\mathbf{W}$ matrix has been obtained by moving $\mathbf{V}_N^{-1}$ along the main diagonal in four blocks, while it has been shifted in three blocks when $N = 560$. As it can be seen, most of the weight for $\widetilde{h}_i$ and $\widehat{h}_i$ is on the nearest observations to the $i-$th one and the true weights given to remote values are nearly zero. Moreover, in the neighborhood of the $i-$th observation, the approximated weights are indistinguishable from the true weights, and the differences between them in the tails are negligible when $N = 560$. As expected, the smaller is N the larger the differences, but even with $N = 350$, the largest discrepancies amount only to approximately 5×10^{-4}. Furthermore, it should be recalled that for a less persistent model and a larger sample size, the differences will become even smaller, because the larger the sample size and the smaller d and σ_η^2, the smaller the true weights assigned to remote observations and the better the approximation.

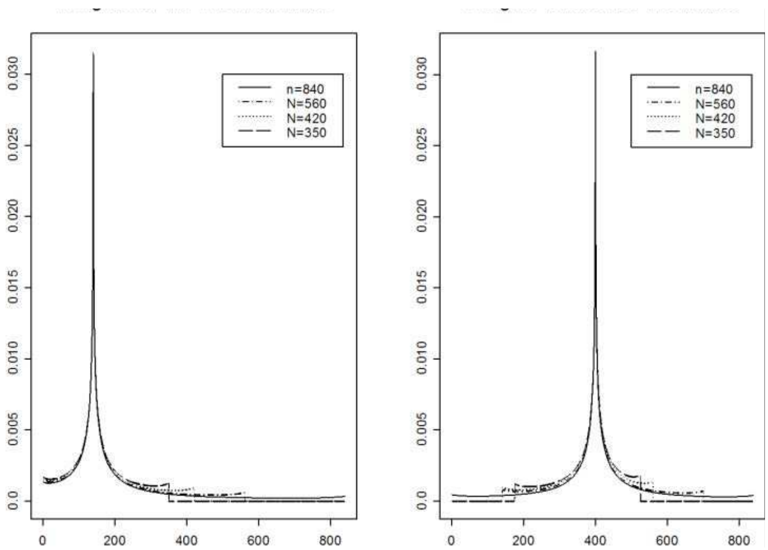

Figure 1: True and approximated weights for the 140-th (left panel) and the 400-th (right panel) observation in a $LMSV$ model with $\{d = 0.45, \sigma_\eta^2 = 0.1, \sigma = 1\}$, sample size $n = 840$ (solid line), and truncation parameter $N = 560$ (dots and dashes), $N = 420$ (dots) and $N = 350$ (dashes).

3. Empirical Application

3.1. Data and Descriptive Statistics

Herein we apply the estimation and smoothing methods discussed in the previous sections to the three series of daily exchange rate data analyzed by Wright (1999b). These series are the Dollar/Pound, Dollar/Mark and Dollar/Yen covering the period 1986-1996 inclusive. The exchange rate returns are constructed as 100 times the first differences of the log of exchange rates. These returns were demeaned. Table 1 reports summary statistics for the unconditional distribution of these series as well as the sample autocorrelations of log-squared returns and the Ljung-Box statistic for uncorrelation in log-squared returns. Figure 2 displays the three series of daily returns and the corresponding correlograms of log-squared returns.

Series	MARK	POUND	YEN
Observations	2696	2680	2680
Median	0.001	0.010	-0.015
St. Dev.	0.289	0.305	0.318
Skewness	-0.085	-0.251	0.346
Kurtosis	5.016	5.679	9.430
Autocorrelations of log-squared returns			
$r_{log2}(1)$	0.176	0.053	0.176
$r_{log2}(2)$	0.054	0.122	0.054
$r_{log2}(5)$	0.039	0.090	0.034
$r_{log2}(10)$	0.087	0.085	0.020
$Q_{log2}(20)$	187.83	342.89	122.06
$Q_{log2}(100)$	328.82	785.24	204.88

$r_{log2}(k)$ denote the kth-order autocorrelation of log-squared returns.
$Q_{log2}(k)$ is the kth Ljung-Box statistic for serial correlation in log-squares.

Table 1: Summary statistics for the return series.

As expected, the three series exhibit excess kurtosis and volatility clustering. Moreover, the Ljung-Box statistics for uncorrelation up to $20-th$ and $100-th$

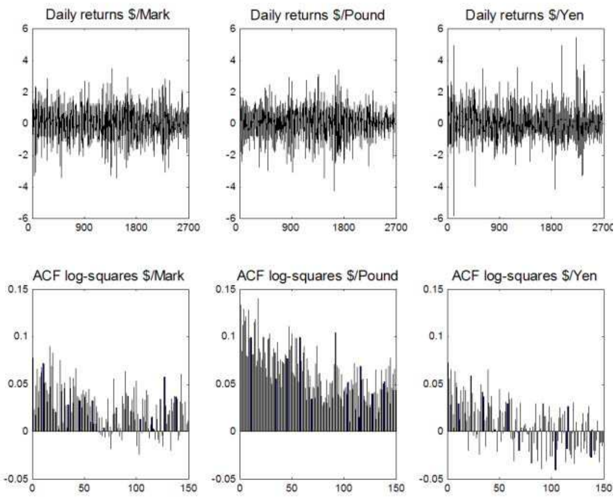

Figure 2: Daily returns (top panels) and correlograms of log-squared returns (bottom panels) for the Dollar/mark, Dollar/Pound and Dollar/Yen exchange rates.

order in the log-squares are all highly significant and so are the selected autocorrelations, except the $5-th$ and the $10-th$ ones of the Yen series. A further look at the correlogram of the log-squares displayed in Figure 2 suggests the presence of long-memory in the Mark and Pound series, characterized by a slow hyperbolic decay of the autocorrelations towards zero, and maybe also in the Yen series, although not so clear. These findings are in agreement with other studies that have also reported evidence on the existence of long memory in the volatility of foreign exchange returns; see, for example, Andersen and Bollerslev (1997), Bollerslev and Wright (2000), Baillie *et al.* (1996) and Hurvich and Ray (2003).

3.2. Tests for Long-Memory in Volatility

Wright (1999b) has reported strong evidence against the presence of a unit root in the volatility of these exchange rate data analyzed here. As he points out, this could indicate that models in which the volatility process is fractionally integrated may provide a better representation of the data. Herein we complement this analysis by testing the hypothesis of $d = 0$ (short memory) against the fractional alternative $d > 0$. In particular, we apply the LM test developed by Lobato and Robinson (1998) and get the following values: 3.16 for the Yen series, 4 for the Mark series and 3.86 for the Pound series. This test rejects the null if the test statistic value falls in the upper tail of the standard Normal distribution. Therefore, it seems that there is strong evidence of long memory in the volatility of the three exchange rate series considered here. In addition, we perform the Wald test for long memory in volatility based on the log-periodogram of the log-squared returns, suggested by Hurvich and Soulier (2002). Under the null hypothesis of $d = 0$, these authors prove that their statistic is asymptotically normal with zero mean and variance $\pi^2/24$. The corresponding test values becomes 3.44, 4.09 and 1.91, for the Mark, Pound and Yen series, respectively, so it confirms the evidence of long memory in the volatility of Mark and Pound, whereas it is less conclusive for the Yen series. Finally, we apply to the log-squared returns the rescaled variance test for long memory proposed by Giraitis *et al.* (2003). Again, this test provides evidence of long memory in the volatility of Mark and Pound series, while it is unable to reject the null of short memory for the Yen series.

3.3. Model Estimation

Given the previous results, we now estimate a *LMSV* model for the three return series by means of the Whittle-type *FDQML* method proposed by Breidt *et al.* (1998) and Harvey (1998). This method is easy to apply and allows us to estimate all the parameters of the model, which is something essential to carry out the smoothing explained in section 2. For the *LMSV* model in (2)-(3), the *FDQML* estimator is obtained by minimizing the following function

$$Q_n(\Theta) = \frac{2\pi}{n} \sum_{j=1}^{[n/2]} \left\{ \log f(\lambda_j; \Theta) + \frac{I_n(\lambda_j)}{f(\lambda_j; \Theta)} \right\},$$

where Θ is the vector of unknown parameters, $[n/2]$ is the integer part of $n/2$, $\lambda_j = 2\pi j/n$, $I_n(\lambda_j)$ is the j-th periodogram ordinate of x_t and $f(\lambda_j;\Theta)$ stands for the spectral density of x_t evaluated at λ_j. In the stationary case ($d < 1/2$), we have

$$f(\lambda_j;\Theta) = \frac{\sigma_\eta^2}{2\pi}\left|\frac{\theta(e^{i\lambda})}{\phi(e^{i\lambda})}\right|[2(1-\cos\lambda_j)]^{-d} + \frac{\sigma_\xi^2}{2\pi}.$$

If the series x_t is not stationary ($1/2 < d < 1$), *FDQML* estimation is carried out for the series $\triangle x_t$, whose spectral density is given by

$$f(\lambda_j;\Theta) = \frac{\sigma_\eta^2}{2\pi}\left|\frac{\theta(e^{i\lambda})}{\phi(e^{i\lambda})}\right|[2(1-\cos\lambda_j)]^{1-d} + \frac{\sigma_\xi^2}{2\pi}[2(1-\cos\lambda_j)].$$

We use several $ARFIMA(p,d,q)$ models to capture both the long memory and the short-run dynamics of the log-volatility h_t. In particular, we estimate $(0,d,0)$, $(1,d,0)$, $(0,d,1)$ and $(1,d,1)$ specifications. To select among the estimated models, we use the Akaike (*AIC*) and Schwarz (*SIC*) information criteria. Both of them select the $(0,d,0)$ specification for the Mark and Pound rate series. However, for the Yen series, the AIC and the SIC provide different results and we choose the more parsimonious $(0,d,0)$ model selected by the SIC. Table 2 (A) reports the estimates of the parameters for the selected *LMSV* models. Interestingly, the estimate of d for the Mark series is in close accordance with previous findings in Baillie *et al.* (1996) using a *FIGARCH* specification, and Harvey (1998) using the *LMSV* model. Moreover, the estimated values of σ_ξ^2 are not far from the one corresponding to a normal distribution for ε_t. In order to compare estimates of the volatility arising from the *LMSV* model with that provided by *GARCH*-type models, we have also fitted to our data a *GARCH* model with long memory that also generates estimates for the log-volatility as in the *LMSV* framework. Such model is the *FIEGARCH* proposed by Bollerslev and Mikkelsen (1996), where the log-volatility is parametrized as an $ARFIMA(p,d,q)$ given by:

$$\phi(B)(1-B)^d \ln\sigma_t^2 = \omega + [1+\psi(B)]g(\varepsilon_{t-1}), \tag{6}$$

where $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$, $\psi(B) = \psi_1 B + \cdots + \psi_q B^q$ and $g(\varepsilon_t) = \theta\varepsilon_t + \gamma[|\varepsilon_t| - E(|\varepsilon_t|)]$. The parameter θ accounts for what is sometimes called 'leverage effect', that is, for $\theta < 0$, the future conditional variance will increase proportionally more following a negative shock than for a positive shock of the same magnitude. Since the *LMSV* model in equations (1)-(2) does not account for this effect[3], and to make the previous results comparable with those from the *FIEGARCH* model, we have estimated the latter with no leverage effect, that is, by fixing $\theta = 0$. This model is estimated by maximum likelihood by assuming ε_t to be $NID(0,1)$. As in the *LMSV* application, several specifications are used for the $ARFIMA(p,d,q)$ process in (6), namely $(0,d,0)$, $(1,d,0)$, $(0,d,1)$ and $(1,d,1)$, and the best model is chosen according to the SIC criteria. Table 2 (B) shows the parameter estimates for the selected models, a $FIEGARCH(0,d,0)$ for Mark and Yen series, and a $FIEGARCH(1,d,0)$ for Pound. The estimate of the fractional difference parameter is statistically very different from both zero and one for the three series analyzed, which suggests that the volatility process is persistent with long memory in all cases.

3.4. Smoothed Volatility Estimates

We obtain, for each series, the estimation of the underlying volatility from the *LMSV* model, using the smoothing algorithm described in section 2 with $N = 1500$. In order to compare these estimated volatilities with those provided by *FIEGARCH* models, Figure 3 displays, for the three series considered, the smoothed estimated volatilities series $\widehat{\sigma}_t$ from the *LMSV* model (left hand side panels) and the *FIEGARCH* model (right hand side panels). This figure shows that, for the three series considered, the smoothed volatility from the *LMSV* model is quite close to the one provided by the *FIEGARCH* model, although the latter seems

[3] A new *LMSV* model that allows for leverage effect has been recently proposed by Ruiz and Veiga (2008) and further analysed by Pérez *et al.* (2009).

Panel A: Whittle *QML* estimates (*LMSV*)			
	MARK	POUND	YEN
	$(0,d,0)$	$(0,d,0)$	$(0,d,0)$
σ	0.805	1.070	0.622
d	0.867	0.716	0.405
σ_η^2	0.019	0.084	0.356
σ_ξ^2	5.093	4.383	4.556
Panel B: Maximum likelihood estimates (*FIEGARCH*)			
	MARK	POUND	YEN
	$(0,d,0)$	$(1,d,0)$	$(0,d,0)$
ω	-0.141 (-6.78)	-0.065 (-7.10)	-0.206 (-13.12)
d	0.711 (19.09)	0.523 (8.89)	0.555 (17.96)
γ	0.184 (6.85)	0.083 (7.14)	0.273 (14.20)
ϕ_1		0.799 (13.48)	

For the *FIEGARCH* model t-statistics are in parenthesis.

Table 2: Estimates of the *LMSV* and *FIEGARCH* models.

to have larger fluctuations than the former, as expected; see Ruiz (1994) and So, Lam and Li (1999) for similar results when comparing short memory *SV* and *GARCH* models.

Moreover, for *LMSV* models (left-hand side graphs), it can be noticed that the estimated volatility of the Mark time series is slightly smoother than that of the Pound series and much smoother than that of the Yen series. This bears out the fact that in *SV* models, the higher the persistence of the signal the smoother the path of the volatility. In contrast, the estimated *FIEGARCH* volatilities (right-hand side graphs) appear to be of roughly the same degree of smoothness for the three series. This is to be expected since modelling a time series directly as a linear function of its past values, as it is the case in *FIEGARCH* models, is not subject to the smoothness constraints imposed in the structural models that underlies the *LMSV* model.

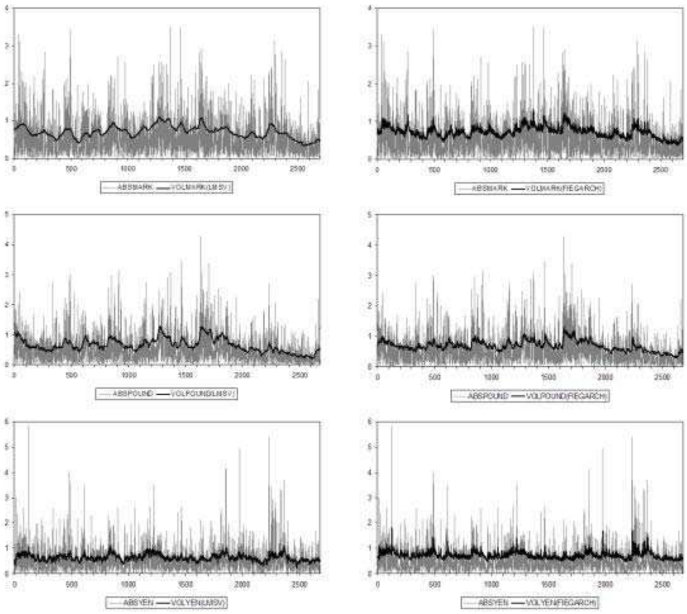

Figure 3: Absolute demeaned returns (thin line) and smoothed volatility (thick line) from LMSV (left hand side panels) and FIEGARCH (right hand side panels) models for Dollar/Mark (top), Dollar/Pound (middle) and Dollar/Yen (bottom) returns.

3.5 Diagnostic Checking and Model Selection

Table 3 reports summary statistics for the standardized returns, $y_t/\widehat{\sigma}_t$, from the two estimated models. For both models, skewness and kurtosis coefficients are smaller than in the original series (compare the values in this table with those in table 1). However, the kurtosis coefficients are still significantly greater than that of the Normal distribution, which means that no model has completely captured the fat-tailed behavior of the returns. This feature is more remarkable in the Yen series and could be due to the presence of outliers. Moreover, notice that in this series, the standardized observations from the *LMSV* model clearly exhibits less kurtosis than those coming from the *FIEGARCH* model. On the other hand, the Box-Ljung statistics for uncorrelation in the log-squared returns up to 20-*th* and 100-*th* order are no longer significant, with p-values greater than 0.05 in all cases. This means that both models successfully deal with the serial correlation observed in the log-squared returns, although *FIEGARCH* models seems to outperform *LMSV* in capturing this feature. Shephard (1996) drew similar conclusions when comparing the empirical performance of short memory *SV* and *GARCH* models. In particular, this author concludes that the success of the *SV* models over the normal-based *GARCH* models is accounted for by its better explanation of the fat-tailed behavior of returns. Also in the short-memory framework, Carnero *et al.* (2004) demonstrate that, although both *GARCH* and *SV* models are able to explain excess kurtosis and significant slowly decaying autocorrelations of squares, the latter is more flexible than the former to represent simultaneously both characteristics. Our empirical results suggest that this property will still hold for heteroscedastic models with long-memory.

	LMSV			*FIEGARCH*		
	Mark	Pound	Yen	Mark	Pound	Yen
Mean	0.023	0.028	0.025	0.002	0.011	0.001
Median	0.046	0.055	-0.026	0.019	0.033	-0.053
St. Dev.	0.999	0.999	0.999	1.002	1.001	1.005
Skewness	0.025	-0.084	0.282	-0.016	-0.201	0.546
Kurtosis	4.519	4.457	5.524	4.565	5.046	7.512
$Q_{log2}(20)$	17.636	22.204	32.658	22.906	17.377	15.837
$Q_{log2}(100)$	101.11	122.73	115.77	93.669	108.21	94.520

Table 3: Summary statistics for the standardized return series.

Finally, to assess the relative performance of the two models we also compare them in terms of the mean squared error (*MSE*) and the mean absolute error (*MAE*) which are defined as $MSE = n^{-1}\sum_{t=1}^{n}(|y_t| - \widehat{\sigma}_t)^2$ and $MAE = n^{-1}\sum_{t=1}^{n}||y_t| - \widehat{\sigma}_t|$, where the absolute demeaned returns, $|y_t|$, act as a proxy for the actual volatility of the returns and $\widehat{\sigma}_t$ denotes the smoothed estimated volatility arising either from the *LMSV* or *FIEGARCH* model. The values of these two measures for each estimated model are given in Table 4. Clearly, both models perform quite similarly, providing very close values of both *MSE* and *MAE*, except for the Yen series, where the *LMSV* model seems to capture the underlying volatility slightly better than the *FIEGARCH*.

	LMSV			*FIEGARCH*		
	Mark	Pound	Yen	Mark	Pound	Yen
MSE	0.250	0.217	0.226	0.271	0.248	0.303
MAE	0.407	0.368	0.357	0.426	0.402	0.435

Table 4: *MSE* and *MAE* for *LMSV* and *FIEGARCH* models

4. Conclusions

In this paper, we propose a computationally efficient method to obtain a smoothed estimation of the unobserved volatility in *LMSV* models, based on the original proposal in Harvey (1998). Estimating the volatility is important from both a theoretical and an empirical point of view since it has a direct implication in the construction of volatility forecasts and in asset pricing and also as an essential tool for diagnostic checking and model selection. The new method hinges on approximating the true $n \times n$ weight matrix of the *MMSLE* of the signal with another $n \times n$ matrix which has zeroes filled in the edges, and whose central elements come from the inverse of an $N \times N$ matrix, N being much smaller than n. We implement our procedure and obtain an estimation of the underlying volatility for the Dollar/Pound, Dollar/Mark and Dollar/Yen series of daily exchange rate data. We also compare our results with those obtained from fitting a *FIEGARCH* model. Our main conclusion is that the proposed algorithm provides an accurate estimation of the underlying volatility that mimics the temporal changes empirically observed in financial time series.

Acknowledgements We wish to dedicate this paper to the memory of our co-author F. Marmol, who passed away after completing a previous version. The second author research is partially supported by Projects VA092A08 and VA027A08 of Junta de Castilla y León and Project SEJ2006-03919 of Ministerio de Educación y Ciencia. The third author gratefully acknowledges financial support from Dirección Xeral de I+D (Xunta de Galicia) under PGIDIT06PXIB201002PR and from Ministerio de Ciencia e Innovación under MTM2008-03010. We would like to thank Esther Ruiz for very helpful comments on a previous draft of this paper.

References

1. Andersen, T.G. and Bollerslev, T. (1997): Heterogeneous Information Arrivals and Return Volatility Dynamics: Uncovering the Long-Run in High Frequency Returns. *Journal of Finance* LII, 975-1005.
2. Andersen, T.G., Bollerslev, T., Diebold, F.X. and Ebens, H. (2001a): The Distribution of Realized Stock Return Volatility. *Journal of Empirical Economics* 61, 43-76.
3. Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. (2001b): The Distribution of Exchange Rate Volatility. *Journal of the American Statistical Association* 96, 42-55.
4. Arteche, J. (2004): Gaussian Semiparametric Estimation in Long Memory in Stochastic Volatility and Signal plus Noise Models. *Journal of Econometrics* 119, 131-154.
5. Baillie, R.T., Bollerslev, T. and Mikkelsen, H.O. (1996): Fractionally Integrated Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 74, 3-30.
6. Bertelli, S. and Caporin, M. (2002): A Note on Calculating Autocovariances of Long-Memory Processes. *Journal of Time Series Analysis* 23, 503-508.
7. Bollerslev, T. and Mikkelsen, H.O. (1996): Modeling and Pricing Long Memory in Stock Market Volatility. *Journal of Econometrics* 73, 151-184.
8. Bollerslev, T. and Wright, J.H. (2000): Semiparametric estimation of Long-Memory Volatility Dependences: The Role of High Frequency Data. *Journal of Econometrics* 98, 81-106.
9. Breidt, F.J., Crato, N. and de Lima, P. (1998): The Detection and Estimation of Long Memory in Stochastic Volatility. *Journal of Econometrics* 83, 325-348.

10. Carnero, M.A., Peña., D. and Ruiz, E. (2004): Persistence and Kurtosis in GARCH and Stochastic Volatility Models. *Journal of Financial Econometrics* 2, 319-342.

11. Chen, W.W., Hurvich, C.M. and Lu, Y. (2006): On the Correlation Matrix of the Discrete Fourier Transform and the Fast Solution of Large Toeplitz Systems for Long-memory Time Series. *Journal of the American Statistical Association* 101 (474), 812-822.

12. Chung, C.F. (1994): A Note on Calculating the Autocovariances of the Fractionally Integrated ARMA Models. *Economics Letters* 45, 293-297.

13. Dacorogna, M.M., Muller, U.A., Nagler, R.J., Olsen, R.B. and Pictet, O.V. (1993): A Geographical Model for the Daily and Weekly Seasonal Volatility in the Foreign Exchange Market. *Journal of International Money and Finance* 12, 413-438.

14. Deo, R. and Hurvich, C. (2001): On the Log-Periodogram Regression Estimator of the Memory Parameter in Long Memory Stochastic Volatility Models. *Econometric Theory* 17, 686-710.

15. Deo, R. and Hurvich, C., and Lu, Y. (2006): Forecasting Realized Volatility Using a Long-Memory Stochastic Volatility Model: Estimation, Prediction and Seasonal Adjustment. *Journal of Econometrics* 131, 29-58.

16. Ding, Z., Granger, C.W.J. and Engle, R.F. (1993): A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance* 1, 83-106.

17. Giriatis, L., Kokosza, P., Leipus, R. and Teyssiere, G. (2003): Rescaled Variance and Related Tests for Long Memory in Volatility and Levels. *Journal of Econometrics* 112, 265-294.

18. Harvey, A.C. (1998): Long Memory in Stochastic Volatility, in J. Knight and S. Satchell (eds.). *Forecasting Volatility in the Financial Markets*, 307-320. London: Butterworth-Heineman.

19. Harvey, A.C. and Shephard, N.G. (1993): Estimation and Testing of Stochastic Volatility Models. *STICERD Econometrics Discussion Paper* EM/93/268, London School of Economics.

20. Hosking, J.R.M. (1981): Fractional Differencing. *Biometrika* 68, 165-176.

21. Hurvich, C.M. and Ray, B. (2003): The Local Whittle Estimator of Long-Memory Stochastic Volatility. *Journal of Financial Econometrics* 1, 445-470.

22. Hurvich, C.M. and Soulier, P. (2002): Testing for Long Memory in Volatility. *Econometric Theory* 18, 1291-1308.

23. Hurvich, C.M., Moulines, E. and Soulier, P. (2005): Estimating Long Memory in Volatility. *Econometrica* 73, 1283-1328.

24. Jensen, M.J. (2004) Semiparametric Bayesian Inference of Long-Memory Stochastic Volatility Models. *Journal of Time Series Analysis* 25, 895-922.

25. Lobato, I. and Robinson, P.M. (1998): A Nonparametric Test for I(0). *Review of Economic Studies* 65, 475-495.

26. Pérez, A. (2002): *Estimation and Indentification in Long Memory Stochastic Volatility Models* (Spanish text), UMI number 3044853, Michigan: UMI Proquest. Also available at http://cervantesvirtual.com/FichaObra.html?Ref=7318.

27. Pérez, A., Ruiz, E. and Veiga H. (2009): A Note on the Properties of Power-Transformed Returns in Long-Memory Stochastic Volatility Models with Leverage Effect. *Computational Statistics and Data Analysis*, forthcoming.

28. Ruiz, E. (1994): Modelos para Series Temporales Heterocesdásticas. *Cuadernos Económicos ICE* 56, 73-108.

29. Ruiz, E. and Veiga, H, (2008): Modelling Long-Memory Volatilities with Leverage Effect_A-LMSV versus FIEGARCH. *Computational Statistics and Data Analysis* 52, 2846-2862..

30. Shephard, N. G. (1996): Statistical Aspects of ARCH and Stochastic Volatility. In *D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielsen (eds.), Time Series Models in Econometrics, Finance and Other Fields*, 1-67. London: Chapman and Hall.

31. So, M.K. (1999): Time Series with Additive Noise. *Biometrika* 86, 474-482.

32. So, M.K. (2002): Bayesian Analysis of Long Memory Stochastic Volatility Models. *Sankhya: The Indian Journal of Statistics*, series B, 64, 1-10.

33. So, M.K., Lam, K. and Li, W.K. (1999): Forecasting Exchange Rate Volatility Using Autoregressive Random Variance Model. *Applied Financial Economics* 94, 583-591.

34. Wright, J.H. (1999a): A New Estimator of the Fractionally Integrated Stochastic Volatility Model. *Economics Letters* 63, 295-303.

35. Wright, J.H. (1999b): Testing for a Unit Root in the Volatility of Asset Returns. *Journal of Applied Econometrics* 14, 309-318.