Stochastic Volatility Models with Long Memory

Clifford M. Hurvich* Philippe Soulier†

1 Introduction

In this contribution we consider models for long memory in volatility. There are a variety of ways to construct such models. Our primary fucus here will be on models in discrete time that contain a latent process for volatility. The most well-known model of this type is the Long-Memory Stochastic Volatility (LMSV) model, proposed independently by Breidt, Crato and de Lima (1998) and Harvey (1998). It is a long-memory generalization of the Stochastic Volatility (SV) model of Taylor (1986). The LMSV model is appropriate for describing series of financial returns at equally-spaced intervals of time. The model implies that returns are a finite-variance Martingale difference sequence, hence uncorrelated, while power transformations of the absolute returns have slowly decaying autocorrelations, in keeping with the empirical findings of Ding, Granger and Engle (1993). We will present the LMSV model, explain its basic properties, and give a survey of existing theoretical results. A variety of generalizations of the model have been considered, and some of these will be briefly discussed, but in order to enhance readability we will focus on a basic form of the model.

An important distinction between ARCH-type models and SV-type models is that the former are observation-driven, giving an expression for the one-step-ahead conditional variance in terms of observables and model parameters, while the latter are driven by a latent (unobserved) process with stands as a proxy for volatility but which does not represent the conditional variance. Thus, for the LMSV model it is necessary to use and develop appropriate techniques in order to carry out basic activities such as forecasting of squared returns, or aggregates of these (i.e., the realized volatility; see, e.g., Andersen, Bollerslev, Diebold and Labys 2001), as well as estimation of parameters.

For simplicity, we will assume that the latent long-memory process is stationary and Gaussian, and is independent of the multiplying shock series (see Equation (2.1) below). We will consider parameter estimation, forecasting, smoothing, as well as semiparametric estimation and hypothesis testing for the long memory parameter. Besides presenting theoretical results, we will also discuss questions of computational efficiency.

*New York University : 44 W. 4'th Street, New York NY 10012, USA
†Université Paris X, 200 avenue de la République, 92001 Nanterre cedex, France

There are several definitions of long memory, which are not equivalent in general (see Taqqu, 2003). For simplicity, we will say here that a weakly stationary process has long memory if its autocovariances $\{c_r\}$ satisfy

$$c_r \sim K_1 r^{2d-1}$$

($K_1 > 0$) as $r \to \infty$, or if its spectral density $f(\omega)$, $\omega \in [-\pi, \pi]$ satisfies

$$f(\omega) \sim K_2 |\omega|^{-2d}$$

($K_2 > 0$) as $\omega \to 0$, where $d \in (0, 1/2)$ is the memory parameter.

2 Basic Properties of the LMSV Model

The LMSV model for a stationary series of returns $\{r_t\}$ takes the form

$$r_t = \exp(Y_t/2) e_t \tag{2.1}$$

where $\{e_t\}$ is a series of i.i.d. shocks with zero mean, and $\{Y_t\}$ is a zero-mean stationary Gaussian long-memory process, independent of $\{e_t\}$, with memory parameter $d \in (0, 1/2)$. Since $\{e_t\}$ is a Martingale difference sequence, so is $\{r_t\}$. It follows that $\{r_t\}$ is a white noise sequence with zero mean. As is the case for most existing volatility models, the LMSV model is nonlinear, in that it cannot be represented as a linear combination of an i.i.d. series.

To study persistence properties of volatility, Ding, Granger and Engle (1993) used power transformations of absolute returns. Using the properties of the lognormal distribution, (see for example Harvey 1998, equation (12.9); cf. Robinson and Zaffaroni 1997, 1998, Robinson 2001) it is possible to derive an explicit expression for the autocorrelations of $\{|r_t|^c\}$ for any positive c such that $\kappa_c = \mathbb{E}(|e_t|^{2c})/\{\mathbb{E}(|e_t|^c)\}^2$ is finite. The expression implies that the $\{|r_t|^c\}$ have long memory with the same memory parameter d, for all such c. It follows that if $\kappa_2 < \infty$ and M is a fixed positive integer, the realized volatility $\{RV_k\}$ given by

$$RV_k = \sum_{t=(k-1)M+1}^{kM} r_t^2$$

has long memory with memory parameter d.

For estimation, it is convenient to work with the logarithms of the squared returns, $\{X_t\} = \{\log r_t^2\}$, which have the signal plus noise representation

$$X_t = \mu + Y_t + \eta_t, \tag{2.2}$$

where $\mu = \mathbb{E}[\log e_t^2]$ and $\{\eta_t\} = \{\log e_t^2 - \mathbb{E}[\log e_t^2]\}$ is an i.i.d. process independent of $\{Y_t\}$. Thus, the log squared returns $\{X_t\}$ are expressed as the sum of the long-memory process $\{Y_t\}$, the signal, and the i.i.d. process $\{\eta_t\}$, the noise, which is independent of the signal. It follows from (2.2) that the autocovariances of $\{X_t\}$ are equal to those of $\{Y_t\}$ for all nonzero lags. Therefore, the autocorrelations of $\{X_t\}$ are proportional to those of $\{Y_t\}$. Furthermore, the spectral density of $\{X_t\}$ is given by

$$f_X(\omega) = f_Y(\omega) + \sigma_\eta^2/(2\pi), \tag{2.3}$$

where $\sigma_\eta^2 = \text{var}(\eta_t)$, assumed to be finite, and hence we have

$$f_X(\omega) \sim K_2|\omega|^{-2d}$$

($K_2 > 0$) as $\omega \to 0$. Thus, the log squared returns $\{X_t\}$ have long memory with memory parameter d.

3 Parametric Estimation

In both Harvey (1998) and Breidt, Crato and de Lima (1998) it is assumed that $\{Y_t\}$ is generated by a finite-parameter model. This model is taken to be the $ARFIMA(0, d, 0)$ model in Harvey (1998) and the $ARFIMA(p, d, q)$ model in Breidt, Crato and de Lima (1998). Given any finite-parameter long-memory specification for $\{Y_t\}$ in the LMSV model, we face the problem of estimating the model parameters based on observations $r_1, \cdots, r_n$. Full maximum likelihood is currently infeasible from a computational point of view since it would involve an n-dimensional integral. Since long-memory models do not have a state-space representation, it is not possible to directly use a variety of techniques that have been successfully implemented for estimation of autoregressive stochastic volatility models (see, e.g., Harvey, Ruiz and Shephard 1994). We consider here two variants of Gaussian quasi maximum likelihood (QML), in the time and frequency domains. Both are based on the log squared returns, $\{X_t\}_{t=1}^n$, and both are presumably inefficient compared to the (infeasible) full Maximum Likelihood estimator.

The time domain Gaussian QML estimator is based on treating the $\{X_t\}$ as if they were Gaussian, even though in general they will not be Gaussian. Then we can write -2 times the log likelihood function as

$$L(\theta) = \log|\Sigma_{x,\theta}| + (x - \mu_x)'\Sigma_{x,\theta}^{-1}(x - \mu_x) \tag{3.1}$$

where $x = (x_1, \cdots, x_n)'$, θ denotes the parameter vector (consisting of the parameters in the model for $\{Y_t\}$ together with σ_η^2), and μ_x, $\Sigma_{x,\theta}$ are, respectively, the expected value of x and the covariance matrix for x under the model θ. Deo (1995) has established the $\sqrt{n}$ consistency and asymptotic normality of the time domain Gaussian QML estimator. Beyond this theoretical result, there are few if any empirical results available on the performance of this estimator, largely due to computational obstacles, i.e., the calculation of the entries of $\Sigma_{x,\theta}$, the determinant $|\Sigma_{x,\theta}|$ and the quadratic form $(x - \mu_x)'\Sigma_{x,\theta}^{-1}(x - \mu_x)$. These obstacles can be surmounted, however.

In fact, $L(\theta)$ may be calculated in $O(n \log^{3/2} n)$ operations, in the case where $\{Y_t\}$ is assumed to obey an $ARFIMA(p, d, q)$ model. This is achieved by using the Fast Fourier Transform (FFT), which is readily available in standard software such as Matlab and Splus. We briefly sketch the approach, described in detail in Chen, Hurvich and Lu (2006). Since $\{X_t\}$ is weakly stationary, the entries of $\Sigma_{x,\theta}$ are constant along the diagonals, i.e., $\Sigma_{x,\theta}$ is a Toeplitz matrix. The entire matrix is determined by the autocovariances of $\{X_t\}$ at lags $0, \cdots, n-1$. However, it is important here to avoid actually computing the full matrix $\Sigma_{x,\theta}$ since this would require at least n^2 operations, resulting in extremely slow performance when n is large, say, in the hundreds or thousands. Since the autocovariances of $\{X_t\}$ are identical to those of $\{Y_t\}$, calculation of the entries of $\Sigma_{x,\theta}$ reduces essentially to the calculation of the autocovariances of an $ARFIMA(p, d, q)$ model. Analytical expressions for these autocovariances were obtained by Sowell (1992). These expressions involve the hypergeometric function. Numerically, the autocovariances can be computed to any desired degree of accuracy in $O(n \log n)$ operations using the algorithm of Bertelli and Caporin (2002). Chen, Hurvich and Lu (2006) present a preconditioned conjugate gradient (PCG) algorithm for computing the quadratic form in (3.1) in $O(n \log^{3/2} n)$ operations. They also present an accurate approximation to the determinant term in (3.1) due to Böttcher and Silbermann (1999) which can be computed in $O(1)$ operations. The PCG method for calculating the likelihood is faster than the $O(n^2)$ that would be required based on the algorithm of Levinson (1946) as advocated by Sowell (1992).

Breidt, Crato and de Lima (1998) proposed to estimate the parameters of the LMSV model from $\{X_t\}$ using the Whittle approximation to the likelihood function. Given data $x_1, \cdots, x_n$, define the periodogram

$$I_j = |\sum_{t=1}^{n} x_t \exp(-i\omega_j t)|^2/(2\pi n) \quad j = 1, \cdots, n-1 \ ,$$

where $\omega_j = 2\pi j/n$ are the Fourier frequencies. Mean correction in the definition above is not necessary since it would not change the values of I_j for $j > 0$. The Whittle approximation for $-2 \log$ likelihood is

$$L_W(\theta) = \sum_{j=1}^{\lfloor (n-1)/2 \rfloor} \{\log f_{X,\theta}(\omega_j) + I_j/f_{X,\theta}(\omega_j)\}$$

where $f_{X,\theta}(\omega_j)$ is the spectral density for X at frequency (ω_j) under the model θ. It is easy to compute $L_W(\theta)$ since $\{I_j\}$ can be computed in $O(n \log n)$ operations using the FFT, and since $f_{X,\theta}$ is the sum of a constant and an $ARFIMA$ spectral density, which has a simple analytical form. Breidt, Crato and de Lima established the consistency of the Whittle estimator. Hosoya (1997) presents results on the $\sqrt{n}$-consistency and asymptotic normality of the Whittle estimator. Perez and Ruiz (2001) have studied the empirical properties of the Whittle estimator for LMSV models.

4 Semiparametric Estimation

In a preliminary econometric analysis, it is often of interest to try to gauge the existence and strength of long memory without imposing a fully parametric model. An easily implemented semiparametric estimator of d is the log-periodogram regression estimator $\hat{d}_{GPH}$ of Geweke and Porter Hudak (1993), obtained as $-1/2$ times the least-squares slope estimate in a linear regression of $\{\log I_j\}_{j=1}^m$ on $\{\log|1-e^{-i\omega_j}|\}_{j=1}^m$, where m tends to ∞ more slowly than n. The $\sqrt{m}$-consistency and asymptotic normality of $\hat{d}_{GPH}$ assuming Gaussianity were obtained by Robinson (1995a) with trimming of low frequencies, and by Hurvich, Deo and Brodsky (1998) without trimming. The latter paper also showed that under suitable regularity conditions the optimal choice of m, minimizing the asymptotic mean squared error, is of order $n^{4/5}$. The regularity conditions were imposed on the short-memory component of the spectral density. For any weakly stationary process with memory parameter d and spectral density f, the short-memory component is defined by $f^*(\lambda) = f(\lambda)/|1-\exp(-i\lambda)|^{-2d}$. The results described above do not apply directly to the estimator $\hat{d}_{GPH}$ based on the log squared returns $\{X_t\}$ in the LMSV model, since in general $\{X_t\}$ will be non-Gaussian (and nonlinear).

For the LMSV model, Deo and Hurvich (2001) established the $\sqrt{m}$-consistency and asymptotic normality of $\hat{d}_{GPH}$ based on $\{X_t\}$ under suitable smoothness conditions on the short-memory part of the spectral density of the signal $\{Y_t\}$. Under these conditions the short-memory part of the spectral density of the log squared returns $\{X_t\}$ behaves like $C+\omega^\beta$ as $\omega \to 0^+$ where $C > 0$ and $\beta = 2d \in (0,1)$. The resulting MSE-optimal choice for m is of order $n^{2\beta/(2\beta+1)}$ and the corresponding mean squared error of $\hat{d}_{GPH}$ is of order $n^{-2\beta/(2\beta+1)}$. Thus, in the LMSV case the optimal rate of convergence of the mean squared error of $\hat{d}_{GPH}$ depends on d and becomes slower as d decreases. This is due to the presence of the noise term in (2.3) which induces a negative bias in $\hat{d}_{GPH}$. For a given value of d, the bias becomes more severe as larger values of m are used. Even for d close to 0.5, this bias is still problematic as the optimal rate of convergence becomes of order $n^{-2/3}$, much slower than the $O(n^{-4/5})$ rate attained in the Gaussian case, under suitable smoothness conditions.

Hurvich and Ray (2003) introduced a local Whittle estimator of d based on log squared returns in the LMSV model. Hurvich, Moulines and Soulier (2005) established theoretical properties of this semiparametric estimator $\hat{d}_{LW}$, which is a generalization of the Gaussian Semiparametric estimator $\hat{d}_{GSE}$ (Künsch 1987; Robinson 1995b). The results of Arteche (2004) imply that in the LMSV context the GSE estimator suffers from a similar limitation as the GPH estimator: in order to attain $\sqrt{m}$-consistency and asymptotic normality the bandwidth m in $\hat{d}_{GSE}$ cannot approach ∞ faster than $n^{2\beta/(2\beta+1)}$, where $\beta = 2d$. The local Whittle estimator avoids this problem by directly accounting for the noise term in (2.3). From (2.3), it follows that as $\omega \to 0^+$ the spectral density of the log squared returns behaves as

$$f_X(\omega) \sim G\omega^{-2d}(1 + h(d,\theta,\omega))$$

where $G = f_Y^*(0)$, $f_Y^*(\omega) = f_Y(\omega)/|\omega|^{-2d}$, and $h(d,\theta,\omega) = \theta\omega^{2d}$ where $\theta = \frac{\sigma_\eta^2}{2\pi f_Y^*(0)}$. We assume here (as did Deo and Hurvich 2001 as well as Hurvich, Deo and Brodsky 1998) that f_Y^* satisfies a local Lipschitz condition of order 2, as would be the case if $\{Y_t\}$ is a stationary invertible $ARFIMA$ or fractional Gaussian noise process.

The local Whittle contrast function, based on the observations $x_1, \dots, x_n$, is defined as

$$\hat{W}_m(\tilde{d}, \tilde{G}, \tilde{\theta}) = \sum_{j=1}^{m} \left\{ \log\left(\tilde{G}\omega_j^{-2\tilde{d}}(1 + h(\tilde{d}, \tilde{\theta}, \omega_j)\right) + \frac{I_j}{\tilde{G}\omega_j^{-2\tilde{d}}(1 + h(\tilde{d}, \tilde{\theta}, \omega_j))} \right\}.$$

Concentrating $\tilde{G}$ out of $\hat{W}_m$ yields the profile likelihood

$$\hat{J}_m(\tilde{d}, \tilde{\theta}) = \log\left(\frac{1}{m} \sum_{j=1}^{m} \frac{\omega_j^{2\tilde{d}} I_j}{1 + h(\tilde{d}, \tilde{\theta}, \omega_j)} \right) + m^{-1} \sum_{j=1}^{m} \log\{\omega_j^{-2\tilde{d}}(1 + h(\tilde{d}, \tilde{\theta}, \omega_j))\}.$$

The local Whittle estimator is any minimand of the empirical contrast function $\hat{J}_m$ over the admissible set $\mathcal{D}_n \times \Theta_n$ (which may depend on the sample size n):

$$(\hat{d}_{LW}, \hat{\theta}_{LW}) = \arg \min_{(\tilde{d}, \tilde{\theta}) \in \mathcal{D}_n \times \Theta_n} \hat{J}_m(\tilde{d}, \tilde{\theta}).$$

Under suitable regularity conditions, Hurvich Moulines and Soulier (2005) show that if $m \to \infty$ faster than $n^{4d/(4d+1)+\delta}$ for some arbitrarily small $\delta > 0$ and if $m^5/n^4 \log^2 m \to 0$, then $m^{1/2}(\hat{d}_{LW} - d)$ is asymptotically Gaussian with zero mean and variance $(1+d)^2/(16\, d^2)$. The first condition on m above is a lower bound, implying that the m for $\hat{d}_{LW}$ must increase faster than the *upper* bound on m needed for $\sqrt{m}(\hat{d}_{GPH} - d)$ to be asymptotically Gaussian with zero mean. Nevertheless, if we allow m to increase sufficiently quickly, the estimator $\hat{d}_{LW}$ attains the rate (to within a logarithmic term) of $O_p(\sqrt{n^{-4/5}})$, essentially the same rate as attained by $\hat{d}_{GPH}$ in the Gaussian case and much faster than the rate attained by either $\hat{d}_{GPH}$ or $\hat{d}_{GSE}$ in the LMSV case.

Accurate finite-sample approximations to the variance of $\hat{d}_{LW}$ are given in Hurvich and Ray (2003).

Sun and Phillips (2003) proposed a nonlinear log-periodogram regression estimator $\hat{d}_{\mathrm{NLP}}$ of d, using Fourier frequencies $1, \dots, m$. They assumed a model of form (2.3) in which the signal is a Gaussian long memory process and the noise is a Gaussian white noise. This rules out most LMSV models, since $\log e_t^2$ is typically non-Gaussian. They partially account for the noise term $\{\eta_t\}$ in (2.3), through a first-order Taylor expansion about zero of the spectral density of the observations. They establish the asymptotic normality of $m^{1/2}(\hat{d}_{\mathrm{NLP}} - d)$ under assumptions including $n^{-4d} m^{4d+1/2} \to \mathrm{Const}$. Thus, $\hat{d}_{\mathrm{NLP}}$, with a variance of order $n^{-4d/(4d+1/2)}$, converges faster than the GPH estimator, but unfortunately still arbitrarily slowly if d is sufficiently close to zero.

Beyond estimation of d, a related problem of interest is semiparametric testing of the null hypothesis $d = 0$ in the LMSV model, i.e., testing for long memory in volatility. Most existing papers on LMSV models make use of the assumption that $d > 0$ so the justification of the hypothesis test requires additional work. The ordinary t-test based on either $\hat{d}_{GPH}$ or $\hat{d}_{GSE}$ was justified in Hurvich and Soulier (2002) and Hurvich, Moulines and Soulier (2005), respectively, without strong restrictions on the bandwidth.

5 Generalizations of the LMSV Model

It is possible to relax the assumption that $\{Y_t\}$ and $\{e_t\}$ are independent in (2.1). A contemporaneous correlation between $\{e_t\}$ and the shocks in the model for $\{Y_t\}$ was allowed for in Hurvich, Moulines and Soulier (2005), as well as Hurvich and Ray (2003), Surgailis and Viano (2002). See Hurvich and Ray (2003) for more details on estimating the leverage effect, known in the (exponential) GARCH models, where the sign of the return in period t affects the conditional variance for period $t + 1$.

It is possible to replace the Gaussianity assumption for $\{Y_t\}$ in the LMSV model by a linearity assumption. This was done in Hurvich, Moulines and Soulier (2005) and Arteche (2004), among others. Surgailis and Viano (2002) showed that under linearity for $\{Y_t\}$ and other weak assumptions, powers of the absolute returns have long memory, with the same memory parameter as $\{Y_t\}$. This result does not require any assumption about the dependence between $\{Y_t\}$ and $\{e_t\}$.

It is also possible to relax the assumption that $d < 1/2$ in the LMSV model. If $d \in (1/2, 1)$ we can say that the volatility is mean reverting but not stationary. Hurvich, Moulines and Soulier (2005) proved consistency of $\hat{d}_{LW}$ for $d \in (0, 1)$ and proved the $\sqrt{m}$-consistency and asymptotic normality of $\hat{d}_{LW}$ for $d \in (0, 3/4)$.

6 Applications of the LMSV Model

We briefly mention some applications of the LMSV and related models. Deo, Hurvich and Lu (2006) consider using the (parametric) LMSV model to construct forecasts of realized volatility. This requires a numerical calculation of the spectral density of the squared returns, and uses the PCG algorithm to determine the coefficients used to construct the forecasts as a linear combination of present and past squared returns.

A long memory stochastic duration (LMSD) model was introduced in Deo, Hsieh and Hurvich (2005) to describe the waiting times (durations) between trades of a financial asset. The LMSD model has the same mathematical form as the LMSV model, except that the multiplying shocks have a distribution with positive support.

Smoothing of the volatility in LMSV models was considered by Harvey (1998), who gave a formula for the minimum mean squared error linear estimator (MMSLE) of $\{Y_t\}_{t=1}^n$ based on the observations $\{X_t\}_{t=1}^n$. Computation of the coefficients in the linear combination involves the solution of a Toeplitz system, and the MMSLE can be efficiently computed using the PCG algorithm. Nevertheless, the MMSLE methodology suffers from some drawbacks, as described (in the LMSD context) in Deo, Hsieh and Hurvich (2005).

References

[1] Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P. Modeling and forecasting realized volatility. *Journal of the American Statistical Association,* **96**, (2001), 42–55

[2] Arteche, J. Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models. *J. Econometrics,* **119**, (2004), 131–154.

[3] Bertelli, S. and M. Caporin. A note on calculating autocovariances of long-memory processes. *J. Time Ser. Anal.*, **23**, (2002), 503–508.

[4] Böttcher, A. and B. Silbermann. Introduction to Large Truncated Toeplitz Matrices. Springer-Verlag, New York, 1999.

[5] Breidt, F. J., Crato, N., and de Lima, P. The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics,* **83,** (1998), 325-348.

[6] Chen, W.W., Hurvich, C.M. and Lu, Y. 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**, (2006), 812-822.

[7] Deo, R. On GMM and QML estimation for the long memory stochastic volatility model. Working paper, 1995.

[8] Deo, R.S. and Hurvich, C.M. On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models. *Econometric Theory,* **17**, (2001), 686-710.

[9] Deo, R.S., Hsieh, M. and Hurvich, C.M. Tracing the source of long memory in volatility. Working paper, 2005.

[10] Deo, R.S., Hurvich, C.M. and Lu, Yi. Forecasting realized volatility using a long memory stochastic volatility model: estimation, prediction and seasonal adjustment. To appear in *Journal of Econometrics* (2006).

[11] Ding, Z., Granger, C. and Engle, R.F. A long memory property of stock market returns and a new model. *Journal of Empirical Finance,* **1**, (1993), 83–106.

[12] Geweke, J. and Porter-Hudak, S. The estimation and application of long memory time series models. *Journal of Time Series Analysis*, **4**, (1983), 221–238.

[13] Harvey, A. C. Long memory in stochastic volatility. In: Knight, J., Satchell, S. (Eds.), Forecasting volatility in financial markets. Butterworth-Heinemann, London, 1998.

[14] Harvey, A.C., Ruiz, E. and Shephard, N. Multivariate Stochastic Volatility Models. *Review of Economic Studies*, **61**, (1994), 247–264.

[15] Hosoya, Y. A limit theory for long-range dependence and statistical inference on related models. *Annals of Statistics,* **25**, (1997), 105–137.

[16] Hurvich, C.M., Deo, R.S. and Brodsky, J. The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. *J. Time Ser. Anal.* **19**, (1998), 19–46.

[17] Hurvich, C.M., Moulines, E. and Soulier, Ph. Estimating long memory in volatility. *Econometrica,* **73**, (2005), 1283–1328

[18] Hurvich, C.M. and B.K. Ray. The local Whittle estimator of long-memory stochastic volatility. *Journal of Financial Econometrics*, **1,** (2003), 445–470.

[19] Hurvich, C.M. and Ph. Soulier. Testing for long memory in volatility. *Econometric Theory* **18** (2002), 1291–1308.

[20] Künsch, H. R. Statistical aspects of self-similar processes. Proceedings of the World Congress of the Bernoulli Society, Tashkent, Vol. 1, (1987), 67–74.

[21] Levinson, N. The Wiener RMS (root mean square) error criterion in filter design and prediction, *J. Math. Phys.*, **25**, (1946), 261–178.

[22] Perez, A. and Ruiz, E. Finite sample properties of a QML estimator of stochastic volatility models with long memory. *Economics Letters*, **70**, (2001), 157–164.

[23] Robinson, P.M. Log-periodogram regression of time series with long range dependence. *Annals of Statistics* **23** (1995a), 1043–1072.

[24] Robinson, P.M. Gaussian semiparametric estimation of long range dependence. *Annals of Statistics* **24** (1995b), 1630–1661.

[25] Robinson, P.M. The memory of stochastic volatility models. *Journal of Econometrics,* **101**, (2001), 195–218.

[26] Robinson, P.M. and Zaffaroni, P. Modelling nonlinearity and long memory in time series. *Fields Institute Communications,* **11**, (1997), 161–170

[27] Robinson, P.M. and Zaffaroni, P. Nonlinear time series with long memory: a model for stochastic volatility. *Journal of Statistical Planning and Inference,* **68**, (1998), 359–371

[28] Sowell, F. Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models, *J. Econometrics*, **53**, (1992), 165–188.

[29] Surgailis, D. and M.C. Viano. Long memory properties and covariance structure of the EGARCH model. *ESAIM* P&S, **6**, (2002), 311–329.

[30] Sun, Y., and Phillips, P.C.B. Nonlinear log-periodogram regression for perturbed fractional processes. *Journal of Econometrics* **115** (2003), 355–389.

[31] Taqqu, M.S. Fractional Brownian Motion and Long-Range Dependence, in *Theory and Applications of Long-Range Dependence*, P. Doukhan, G. Oppenheim, M.S. Taqqu, Eds. (2003). Berlin: Birkhauser

[32] Taylor, S.J. Modelling Financial Time Series. Wiley, New York, 1986.