DOCUMENTOS DE TRABAJO

BILTOKI

D.T. 2002.02

Gaussian Semiparametric Estimation in Long Memory
in Stochastic Volatility and Signal plus Noise Models

Josu Arteche

Departamentos de
Econometría y Estadística (E.A. III),
Fundamentos del Análisis Económico,
Hacienda y Sistemas Fiscales (E.A. II)
e Instituto de Economía Pública

eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

Facultad de Ciencias Económicas.
Avda. Lehendakari Aguirre, 83
48015 BILBAO.

Documento de Trabajo BILTOKI DT2002.02
Editado por los Departamentos de Economía Aplicada II (Hacienda), Economía Aplicada III (Econometría y Estadística), Fundamentos del Análisis Económico (Teoría Económica) e Instituto de Economía Pública de la Universidad del País Vasco.
Depósito Legal No.: BI-1260-02
ISSN: 1134-8984

Gaussian Semiparametric Estimation in Long Memory in Stochastic Volatility and Signal Plus Noise Models

Josu Arteche[*†]
Department of Econometrics and Statistics
University of the Basque Country (UPV-EHU)
48015 Bilbao
Spain

8 April 2002

Abstract

This paper considers the persistence found in the volatility of many financial time series by means of a local Long Memory in Stochastic Volatility model and analyzes the performance of the Gaussian semiparametric or local Whittle estimator of the memory parameter in a long memory signal plus noise model which includes the Long Memory in Stochastic Volatility as a particular case. It is proved that this estimate preserves the consistency and asymptotic normality encountered in observable long memory series and under milder conditions it is more efficient than the estimator based on a log-periodogram regression. Although the asymptotic properties do not depend on the signal-to-noise ratio the finite sample performance rely upon this magnitude and an appropriate choice of the bandwidth is important to minimize the influence of the added noise. I analyze the effect of the bandwidth via Monte Carlo. An application to a Spanish stock index is finally included.

JEL classification: C13; C22.
Keywords: Long memory; Stochastic Volatility; Semiparametric estimation; Frequency domain.

[*]Corresponding author: Dpt. Economía Aplicada III (Econometría y Estadística); University of the Basque Country (UPV-EHU); Avda. Lehendakari Aguirre 83; 48015 Bilbao; Spain; Tl: 34 94 601 3852; Fax: 34 94 601 3754; E-mail: ja@alcib.bs.ehu.es.

[†]Research supported by grant 9/UPV 00038.321-13503/2001 and Basque Government grant PI-1999-70 of the Departamento de Educación, Universidades e Investigación. I thank the participants in a seminar at the LSE for helpful and constructive comments.

1 Introduction

Uncertainty is nowadays crucial in the modelling of financial time series. In most asset pricing theories the uncertainty associated with the price of the asset is an important factor in the determination of the risk premium. This uncertainty has been usually related with conditional variances which are changing through time trying to accommodate the distinct behaviour of many financial time series. In particular many financial series such as asset returns do not have a marked structure in their autocorrelations (in accord with the efficient market hypothesis) but show a strong persistence in the autocorrelations of some transforms such as squares or other powers of absolute values.

The seminal works of Engle (1982) and Taylor (1986) established the basis of the two prominent tendencies used in the modelling of conditional heteroskedasticity. The Autoregressive Conditional Heteroskedastic (ARCH) models of Engle (1982) and succeeding extensions consider the conditional variance an exact function of the squares of past observations. The second tendency is related with the stochastic volatility (SV) models (Taylor, 1986, Harvey, Ruiz and Shephard, 1994) in which the volatility component is generated by a stochastic process so that both the mean and volatility equations have separate error terms.

A large body of research suggests that the volatility of many financial time series displays strong persistence which cannot be modelled by standard ARCH or SV models. One of the first attempts to model such a behaviour is the Integrated Generalized ARCH (IGARCH) model proposed by Engle and Bollerslev (1986). Although the IGARCH class of models bears much resemblance to the ARIMA for conditional first moments, the analogy is far from complete. For instance, although not covariance stationary, the IGARCH is strictly stationary and ergodic (Nelson, 1990). Moreover the autocorrelation function of the squares is not constant but decreases exponentially (Ding and Granger, 1996). The existence of long memory in powers of the absolute value of the returns of various asset prices (Ding, Granger and Engle, 1993) opened a new branch of research which includes this persistence in the modelling of the conditional variance. Baillie, Bollerslev and Mikkelsen (1996), Bollerslev and Mikkelsen (1996) and Ding and Granger (1996) proposed the Fractionally IGARCH (FIGARCH) which are parametric restrictions of the more general model in Robinson (1991).

More recently Harvey (1998) and Breidt, Crato and de Lima (1998) proposed the Long Memory in Stochastic Volatility (LMSV) processes which model the log-volatility term as a Fractional ARIMA.

None of the cited works has paid attention to the possible existence of a stochastic cyclical behaviour in the volatility of the series which is a relevant characteristic of intraday data (Andersen and Bollerslev, 1997). For example it is widely accepted a distinct U or inverted J shape in the volatility of intraday stock returns over the trading day. If some cyclical pattern is found seasonal dummies are usually employed. I allow for a time evolving seasonality and consider the possibility of a persistent stochastic seasonality in the volatility in the form of Seasonal or Cyclical Long Memory (SCLM) as proposed by Arteche and Robinson (1999). A SCLM volatility is more adequately modelled as a SV process rather than extending FIGARCH models, and this is the practice I adopt in this paper. SV models moreover has the advantage of being the natural discrete time analogue of the continuous time models used in option pricing. In addition their statistical properties are easier to derive than the FIGARCH case. They have the disadvantage of a difficult evaluation of the exact likelihood. However, after linearizing the model takes the form of a signal plus noise which simplifies the analysis. Based on this, Harvey (1998) and Breidt et al. (1998) suggest a Whittle type spectral quasi maximum likelihood estimation which has a very easy implementation. Recently, and independently of this work, Deo and Hurvich (2001) considered semiparametric estimation of the memory parameter in a LMSV model and proposed the Geweke and Porter-Hudak (1983) estimator based on a log-periodogram regression. Its asymptotic normality relies on the Gaussianity of the signal or volatility process, which seems quite restrictive in the series this models focus on. I prefer the more efficient Gaussian semiparametric or local Whittle estimator which relax this restriction and consequently seems more adequate for the estimation of the memory parameter of the volatility in financial time series. Although the main motivation of this paper is the estimation of the persistence of the volatility in LMSV models I prove the validity of the Gaussian semiparametric estimator not only in these models but in a more general signal plus noise setup.

The structure of the paper is as follows. Section 2 considers different alternatives for the modelling of long memory volatility and describes the local long memory in stochastic volatility models. Unlike Harvey (1998) and Breidt et al. (1998) only a partial spectral behaviour is imposed which allows for a wider range of processes. The persistence of an unexpected shock on the volatility of these series is measured by a single memory parameter (in the standard long memory case) or a finite number of parameters (under seasonal long memory). Section 3 focuses on the Gaussian semiparametric or local Whittle estimation of the memory parameters in a signal plus noise model which includes the LMSV as a particular case. Deo and Hurvich (2001) argue that this estimator can not be applied under LMSV. This section shows that under milder conditions than those required by Deo and Hurvich for the log-periodogram regression, the Gaussian semiparametric estimator is consistent and asymptotically normal with lower asymptotic variance. Section 4 considers the finite sample behaviour paying especial attention to the relevance of the bandwidth. I compare an appropriate version of the approximate optimal mean square error bandwidth of Henry and Robinson (1996) with that which minimizes the Monte Carlo mean square error. For practical purposes, and due to the infeasibility of the bandwidth of Henry and Robinson, a data-driven bandwidth is also proposed. Section 5 shows an application to a series of returns of the Spanish stock index Ibex 35. Finally Section 6 concludes and proposes further extensions. Technical details are placed in the Appendix.

2 Long memory in volatility

Consider the discrete time real valued process

$$x_t = \sigma \sigma_t \varepsilon_t \tag{1}$$

where $E_{t-1}\varepsilon_t = 0$ and $Var_{t-1}(\varepsilon_t) = 1$, where the subindex $t-1$ means that the expectations are conditional on past information, and σ is a positive constant. If σ_t is an exact function of information at time $t-1$, x_t has mean zero and is serially uncorrelated with conditional variance $\sigma^2\sigma_t^2$. Robinson (1991) considered σ_t^2 a function of past squared observations

$$\sigma_t^2 = \mu + \sum_{j=1}^{\infty} \theta_j x_{t-j}^2.$$

with the θ_j allowing for long range dependence. A parametric restriction is

$$\sigma_t^2 = \alpha_0 + \beta(L)\sigma_t^2 + [1 - \beta(L) - \pi(d, L)\alpha(L)]x_t^2. \quad (2)$$

where $\beta(L) = \beta_1 L + ... + \beta_p L^p$, $\alpha(L) = 1 + \alpha_1 L + ... + \alpha_q L^q$, and α_0 is a positive constant. If $\pi(d, L) = (1 - L)^d$ then (2) represents the FIGARCH of Baillie et al. (1996) and Bollerslev and Mikkelsen (1996), if $\pi(d, L) \equiv 1$ (2) is the GARCH of Bollerslev (1986) and if in addition $\beta(L) \equiv 0$ then (2) is the ARCH model of Engle (1982). Considering the serially uncorrelated mean zero process $z_t = x_t^2 - \sigma_t^2$, (2) can be written

$$\pi(d, L)\alpha(L)x_t^2 = \alpha_0 + (1 - \beta(L))z_t$$

where the possibility of long memory in x_t^2 is more clearly established. Since $\pi(d, 1) = 0$ for the FIGARCH, the variance of x_t is infinite for $\alpha_0 > 0$, so that x_t is not covariance stationary. However, as in the IGARCH case, the FIGARCH process is strictly stationary and ergodic for $d \leq 1$.

For this process to be well defined and the conditional variance to be positive almost surely all the θ_j must be nonnegative. General conditions on the initial parameters are hard to establish especially with SCLM in x_t^2 of the form $\pi(d, L) = (1-L)^{d_0}\{\prod_{k=1}^{h-1}(1-2L\cos\omega_k + L^2)^{d_k}\}(1 + L)^{d_h}$ with some $d_k \neq 0$ for $k = 1, ..., h$, which produces alternation of positive and negative coefficients in its expansion (Giraitis and Leipus, 1995). Two alternatives that guarantee positiveness of σ_t^2 are the modelling of the persistence in $\log \sigma_t^2$ as suggested by Nelson (1991) and Bollerslev and Mikkelsen (1996) in the standard long memory case and the modelling of σ_t in (1) instead of σ_t^2 as in Robinson (1991) and Robinson and Zaffaroni (1997). These models have the convenient features that the coefficients are not restricted to be positive. However the statistical analysis and estimation of this class of models is quite burdensome even in the short memory case.

In this paper I adopt a different approach based on SV models in the lines of those discussed by Harvey (1998) and Breidt et al. (1998) for parametric LMSV and Deo and Hurvich (2001) in a semiparametric setup. The local LMSV is defined by (1) and

$$\sigma_t = \exp(v_t/2) \quad (3)$$

where the ε_t are iid with mean zero and variance 1 and v_t is a stationary long memory (LM) process whose spectral density satisfies as $\lambda \to 0$

$$f_v(\omega \pm \lambda) \sim C\lambda^{-2d} \tag{4}$$

for some frequency ω, $|d| < 1/2$, and $0 < C < \infty$. The standard local LMSV satisfies (4) for $\omega = 0$. Recently Arteche and Robinson (2000) have generalized (4) allowing a different spectral behaviour at frequencies just after and just before some $\omega \neq 0, \mathrm{mod}(\pi)$ and proposed the Seasonal or Cyclical Asymmetric Long Memory (SCALM) process characterized by a spectral density satisfying as $\lambda \to 0^+$

$$\begin{aligned} f_v(\omega + \lambda) &\sim C_1\lambda^{-2d_1} \\ f_v(\omega - \lambda) &\sim C_2\lambda^{-2d_2} \end{aligned} \tag{5}$$

where C_1, d_1 can be different from C_2, d_2. Then (4) is a restriction of (5) that imposes $C_1 = C_2 = C$ and $d_1 = d_2 = d$. For simplicity of exposition I focus hereafter on symmetric LM acknowledging the possibility of SCALM.

If v_t is a stationary process distributed independently of ε_t then x_t is both covariance and strictly stationary (in fact $\gamma_x(h) = Ex_tx_{t-h} = 0$ for $h \neq 0$). If in addition v_t is Gaussian

$$Ex_t^r = \varkappa_r\sigma^r \exp\left(\frac{r^2}{8}\gamma_v(0)\right) \qquad r = 0, 1, ...,$$

where $\varkappa_r = E\varepsilon_t^r$ and $\gamma_v(h) = Ev_tv_{t+h}$ is the autocovariance of v_t ($\gamma_v(0)$ is thus its variance). In particular if ε_t is $N(0,1)$ we have $Ex_t = Ex_t^3 = 0$, $Var(x_t) = \sigma^2\exp(\gamma_v(0)/2)$ and $Ex_t^4 = 3\sigma^4\exp(2\gamma_v(0))$ so that x_t is white noise with kurtosis $3\exp(\gamma_v(0))$. The excess kurtosis in these models can thus be caused by a large variance of v_t or by a thick tailed distribution of ε_t (e.g. a t distribution). Leverage effects in the form of an asymmetric response of the volatility to positive and negative shocks can be introduced by an appropriate correlation between mean and volatility equation errors.

Taking logs of the squares of x_t in (1) with (3) we have

$$y_t = \log x_t^2 = \mu + v_t + \xi_t \tag{6}$$

where $\mu = \log\sigma^2 + E\log\varepsilon_t^2$ and $\xi_t = \log\varepsilon_t^2 - E\log\varepsilon_t^2$ is iid with zero mean and variance σ_ξ^2. For example if $\varepsilon_t \sim N(0,1)$ then ξ_t is a centered $\log\chi_1^2$ variable with $E\log\varepsilon_t^2 = -1.27$ and $\sigma_\xi^2 = \pi^2/2$. Apart from the constant μ, y_t takes the form of a signal plus noise model where the signal is a LM process uncorrelated with the noise which in this case is (non-Gaussian) iid. The autocovariance function of y_t is

$$\gamma_y(h) = Ey_t y_{t+h} = \gamma_v(h) + \sigma_\xi^2 I_{h=0} \tag{7}$$

where $I_{h=0} = 1$ if $h = 0$ and 0 otherwise[1] and the autocovariances of y_t coincide with those of the signal v_t. Although the local specification in (4) does not restrict the parametric form of v_t at frequencies far from ω, it is interesting for practical and illustrative purposes to briefly describe some relevant parametric LM models. The GARMA (0,0) process is defined by $(1 - 2L\cos\omega + L^2)^d v_t = u_t$ with u_t white noise with variance σ_u^2, and its autocovariance of order h is

$$\gamma_v(h) = \frac{\sigma_u^2}{2\sqrt{\pi}}\Gamma(1-2d)\{2\sin\omega\}^{0.5-2d}\{P_{h-0.5}^{2d-0.5}(\cos\omega) + (-1)^h P_{h-0.5}^{2d-0.5}(-\cos\omega)\} \tag{8}$$

where $\Gamma(z)$ and $P_a^b(z)$ are the gamma and associated Legendre functions respectively (Chung, 1996). If $\omega = 0$ (8) reduces to

$$\gamma_v(h) = \sigma_u^2 \frac{(-1)^h \Gamma(1-4d)}{\Gamma(h-2d+1)\Gamma(1-h-2d)}$$

(Hosking, 1981) and for $\omega = \pi/2$

$$\gamma_v(h) = \frac{\sigma_u^2}{2}\frac{\Gamma(1-2d)}{\Gamma(1-d-\frac{h}{2})\Gamma(1-d+\frac{h}{2})}\{1+(-1)^h\}$$

(Arteche and Robinson, 2000). Correspondingly the spectral density of y_t is

$$f_y(\lambda) = f_v(\lambda) + \frac{\sigma_\xi^2}{2\pi} \qquad \text{for } -\pi \le \lambda \le \pi \tag{9}$$

where if v_t is a GARMA(0,0) process

$$f_v(\lambda) = \frac{\sigma_u^2}{2\pi} 2^{-2d} |\cos\lambda - \cos\omega|^{-2d} \tag{10}$$

and expressions for $\omega = 0$ and $\omega = \pi/2$ are easily obtained from (10). Perhaps the more general parametric LM model is that used by Chan and Wei (1988), Robinson (1994), Chan and Terrin (1995) and Giraitis and Leipus (1995) which has the form

[1] As shown in Harvey et al. (1994) independence between v_t and ε_t is not strictly needed for this result to hold.

$$(1-L)^{d_0}\{\prod_{k=1}^{h-1}(1-2L\cos\omega_k+L^2)^{d_k}\}(1+L)^{d_h}v_t=u_t \tag{11}$$

where the ω_k can be any frequency between 0 and π and u_t is a short memory process with continuous and positive spectrum. (11) allows for different cyclical behaviour across different frequencies. If only $d_0 \neq 0$ then (11) is the fractional ARIMA of Granger and Joyeux (1980) and Hosking (1981). If there is only one $d_k \neq 0$ for $k \neq 0$ then (11) represents the Gegenbauer or GARMA process of Gray, Zhang and Woodward (1989) and Chung (1996). The spectral density of v_t satisfies (4) for $\omega = 0, \omega_1, ..., \omega_{h-1}$ and π and the asymptotic behaviour of the autocovariances is

$$\gamma_v(j) = Ev_tv_{t+j} = \int_{-\pi}^{\pi} f_v(\lambda)\cos(j\lambda)\mathrm{d}\lambda \sim K\sum_{k=0}^{h} j^{2d_k-1}\cos(j\omega_k) \tag{12}$$

as $j \to \infty$ for some finite constant K (Giraitis and Leipus, 1995 and Arteche and Robinson, 2000). Thus the autocovariances show not only the slow hyperbolic decay typical of long memory but, for $\omega_k \neq 0$, also oscillations of amplitude depending on the ω_k. There exist other LM processes non nested in (11) which similarly satisfy (4). For example the seasonal fractional noise of Jonas (1983) or Carlin and Dempster (1989) is characterized by lag-j autocorrelations

$$\rho_x(j) = \frac{1}{2}\left(\left|\frac{j}{S}+1\right|^{2d+1} - 2\left|\frac{j}{S}\right|^{2d+1} + \left|\frac{j}{S}-1\right|^{2d+1}\right)$$

and a spectral density satisfying (4) at $\omega_s = 2\pi s/S$ for $s = 0, 1, ..., [S/2]$.

3 Gaussian semiparametric estimation

Maximum likelihood estimation and statistical inference are hard to implement in stochastic volatility models even under short memory log-volatility due to the existence of separate errors in the mean and log-volatility equations. A less efficient technique, the method of moments, has been suggested by Melino and Turnbull (1990) and Taylor (1986). Its efficiency reduces substantially under log-volatility persistence of the form of a nearly non-stationary AR (Jacquier, Polson and Rossi, 1994). Harvey et al. (1994) proposed a quasi-maximum likelihood method based on the Kalman filter in short memory stochastic volatility models. For long memory the Kalman filter requires truncation in the AR expansion of v_t which can lead to a loss of relevant information under strong persistence. Jacquier et al. (1994) developed a Bayesian technique based on Markov chain simulation in an autoregressive stochastic volatility model. This method is computationally intensive and it is not known how it performs with long memory stochastic volatility.

The form of the spectral density in (9) led Harvey (1998) and Breidt et al. (1998) to propose spectral estimation strategies which are computationally simple. They are based on the frequency domain quasi-maximum likelihood estimation proposed by Whittle (1953) and analysed in a long memory context by Fox and Taqu (1986), Giraitis and Surgailis (1990), Heyde and Gay (1993) and Dahlhaus (1994). Cheung and Diebold (1994) found that this estimate has a similar finite sample efficiency to exact maximum likelihood, especially when the mean of the process is unknown since it does not require mean correction. The estimates are obtained by minimizing

$$L(\beta) = \frac{1}{2\pi n} \sum \left\{ \log f(\lambda_j) + \frac{I_y(\lambda_j)}{f(\lambda_j)} \right\}, \tag{13}$$

where $f(\lambda)$ is the spectral density of the series under analysis, the sum is for $j = 1, ..., [n/2]$ for n the sample size and $[\cdot]$ denoting the integer part, except those j where $f(\lambda_j) = 0$ or ∞; $\lambda_j = 2\pi j/n$ are Fourier frequencies and

$$I_y(\lambda) = |W_y(\lambda)|^2, \qquad W_y(\lambda) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} y_t \exp(-it\lambda)$$

are the periodogram and discrete Fourier transform of y_t, $t = 1, 2, ..., n$, at frequency λ. The properties of this estimate depends on a correct specification of the spectral density. To avoid the possible inconsistency of the estimates of the memory parameters due to spectral misspecification at frequencies far from those of interest, Robinson (1995b) in the standard long memory case and Arteche and Robinson (2000) under SCLM adopted an idea first suggested by Kunsch (1987) and proposed a narrow frequency band version of (13) such that the estimates of C and d, $\tilde{C}$ and $\tilde{d}$, are obtained by minimizing

$$Q(C, d) = \sum{}' \left\{ \log C|\lambda_j|^{-2d} + \frac{|\lambda_j|^{2d}}{C} I_{yj} \right\} \tag{14}$$

where $I_{yj} = I_y(\omega + \lambda_j)$, $\sum'$ runs for $j = \pm 1, ..., \pm m$ if $\omega \neq 0, \pi$ and due to the symmetry of the periodogram for $j = 1, ..., m$ if $\omega = 0, \pi$, and $m < [n/2]$ is the *bandwidth* parameter which goes to infinity more slowly than n. SCALM can be allowed in Q but a trimming of frequencies close to ω is then required (Arteche and Robinson, 2000). Concentrating C out of the objective function, the estimate $\tilde{d}$ is the argument that minimizes

$$R(d) = \log \tilde{C}(d) - \frac{2d}{\delta_\omega m} \sum{}' \log |\lambda_j| \tag{15}$$

over a compact set $\Theta = [\Delta_1, \Delta_2]$, where $\delta_\omega = 1$ if $\omega = 0, \pi$ and 2 otherwise and

$$\tilde{C}(d) = \frac{1}{\delta_\omega m} \sum{}' |\lambda_j|^{2d} I_{yj}.$$

Only (4) is imposed on the form of the spectral density. Far from ω $f(\lambda)$ can be bounded, unbounded or zero and only integrability for covariance stationarity is assumed. Under some mild conditions (for example Gaussianity is not required) these estimates are consistent and have a normal asymptotic distribution (Robinson, 1995b and Arteche and Robinson, 2000).

In the signal plus noise and stochastic volatility models the spectral density of y_t ($f_y(\lambda)$ in (9)) inherits the asymptotic behaviour of $f_v(\lambda)$ if the memory parameter of interest is positive so that f_y diverges at ω as in (4). In the negative memory case $f_y(\omega \pm \lambda) \sim \sigma_\xi^2/2\pi + C\lambda^{-2d}$ as $\lambda \to 0$ so that (4) does not hold and the minimization of (15) is not directly applicable. Thus I focus on the (empirically more relevant) positive and stationary memory case $1/2 > d > 0$.

The linearized local LMSV model in (6) is a particular case of a signal plus noise so that the estimation of the memory parameter of the volatility can be generalized to the estimation of the memory parameter of the signal in such models. To permit this generalization I do not restrict ξ_t to be iid but allow weak dependence as stated in Assumption A.1. In the results derived hereafter independence between v_t and ξ_t is also imposed.

A.1: $\xi_t = \sum_{j=0}^{\infty} \theta_j z_{t-j}$, $z_t \sim iid(0, \sigma_\xi^2)$, $Ez_t^4 = \mu_4 < \infty$ and $\sum_{j=0}^{\infty} |\theta_j| j^{1/2} < \infty$.

Assumption A.1 guarantees that the spectral density of ξ_t is positive and bounded away from zero and include the white noise as a particular case. Under a dependent ξ_t, x_t in (1) is no longer a martingale difference as the local LMSV assumes, but may show a persistent dependence due to the v_t term.

The assumption A.2 imposes an upper bound in the bandwidth.

A.2:

$$\frac{1}{m} + \frac{m^{1+4d}}{n^{4d}}(\log m)^2 \to 0 \qquad \text{as } n \to \infty.$$

Assumption A.2 requires the proportion of frequencies used in the estimation go to zero faster the lower d is. If $m \sim n^{\theta}$ then A.2 entails $\theta < 2d/(1+2d)$. Assumption A.2 is coherent with Assumption A.4' in Robinson (1995b) since in the signal plus noise model in (6) the spectral density of y_t is

$$f_y(\omega \pm \lambda) \sim C\lambda^{-2d}(1 + O(\lambda^{\min(\alpha,2d)}))$$

under A.1, where α represents the degree of smoothness of the spectral density of v_t in Assumption A.1' in the Appendix. If $\alpha \geq 2d$ the degree of spectral smoothness of y_t is $2d$ and A.2 is equivalent to Assumption A.4' in Robinson (1995b). Not surprisingly A.2 is the condition imposed by Deo and Hurvich (2001) to obtain the asymptotic normality of the Geweke and Porter-Hudak estimator. However the Gaussian semiparametric estimator requires milder assumptions on the noise since in Deo and Hurvich (2001) ξ_t is white noise with finite eighth moment whereas A.1 weakens significantly that condition. In addition Gaussianity of v_t is not needed for consistency and asymptotic normality of $\tilde{d}$.

Theorem 1 *Let y_t be defined in (6) and v_t be a LM process with $0.5 > d > 0$. Under assumption A1'-A4' in the Appendix and if ξ_t satisfies assumption A.1 then $\tilde{d} \xrightarrow{p} d$ as $n \to \infty$.*

Assumption A.1'-A.4' are those required for consistency of a fully observable long memory process as imposed in Robinson (1995b) and Arteche and Robinson (2000) for SCLM. The conditions in v_t are a local specification of f_v around ω and linearity with finite second conditional moments of the innovations. For a symmetric SCLM v_t the condition on m is $m^{-1} + mn^{-1} \to 0$ as $n \to \infty$. For SCALM a trimming of the frequencies closest to ω is needed as explained in Arteche and Robinson (2000).

Breidt et al. (1998) showed the consistency of the parametric Whittle estimation in standard LMSV models but no asymptotic distribution was provided. However, the local Whittle estimator is asymptotically normal not only in local LMSV but also in more general signal plus noise models, as stated in Theorem 2.

Theorem 2 *Let* y_t *be defined in (6) and* v_t *be a LM process with* $0.5 > d > 0$. *Under Assumptions A.1', A.5'-A.7' in the Appendix, A.1 and A.2* $\sqrt{4\delta_\omega m}(\tilde{d} - d) \xrightarrow{d} N(0,1)$ *as* $n \to \infty$.

Again the assumptions in the Appendix correspond to those required for the asymptotic normality of $\tilde{d}$ in Robinson (1995b) and Arteche and Robinson (2000) for an observable long memory series and no further comment is needed. The proofs of both theorems are based on Robinson (1995b) and Arteche and Robinson (2000) and are briefly described in the Appendix. The additional conditions required in v_t are basically finite fourth conditional moments of the innovations and

$$\frac{1}{m} + \frac{m^{1+2\alpha}}{n^{2\alpha}}(\log m)^2 \to 0 \qquad \text{as } n \to \infty. \tag{16}$$

Deo and Hurvich (2001) proved the asymptotic normality of the estimator based on the log periodogram regression in a more restrictive setup. In particular they assume $\omega = 0$ and

$$f_{v1}(\lambda) = \left| 2\sin\left(\frac{\lambda}{2}\right)\right|^{-2d} g(\lambda)$$

such that $g'(0) = 0$ and the second derivative of $g()$ is bounded around zero. Therefore f_{v1} satisfies Assumption A.1' for $\omega = 0$ and $\alpha = 2$. Since $d < 1/2$ assumption A.2 implies (16) and the asymptotic normality of $\tilde{d}$ is established in this case under the same restrictions on m imposed by Deo and Hurvich (2001). The advantages of $\tilde{d}$ rest on the milder restrictions on v_t and ξ_t and its greater efficiency with a lower asymptotic variance.

Remark 1: Similar results, although with a more restrictive bandwidth m, are obtained for a LM ξ_t. The Appendix shows that if $f_\xi(\omega \pm \lambda) \sim C_\zeta \lambda^{-2c}$ as $\lambda \to 0$ for C_ζ a positive constant and $c < d$, the asymptotic distribution of $\tilde{d}$ in theorem 2 remains valid as long as $m^{1+4(d-c)}/n^{4(d-c)}\log^2 m \to 0$ as $n \to \infty$.

Remark 2: Giraitis, Robinson and Samarov (1997) obtained an optimal rate of convergence of semiparametric estimators of the memory parameter based on the local smoothness of the spectral density at the origin. An equivalent rate is expected at any other frequency ω as in Assumption A.1'. Under Assumption A.1' the signal pus noise model in (6) satisfies

$$f_y(\omega \pm \lambda) = C\lambda^{-2d}(1 + O(\lambda^{\alpha^*})) \text{ as } \lambda \to 0^+ \tag{17}$$

where $\alpha^* = min(\alpha, 2d)$. For these models the optimal rate of convergence as defined in Giraitis et al. (1997) is n^{-r} for $r = \alpha^*/(2\alpha^* + 1)$ and in view of Assumptions A.2 and A.7' this rate is possibly achieved by the Gaussian semiparametric estimator, although a rigorous proof is needed.

Remark 3: Statistical inference on the memory parameters of the signal can be readily performed in this local Whittle quasi-maximum likelihood context. Tests of spectral symmetry $d_1 = d_2$ against SCALM have been proposed by Arteche and Robinson (2000). Similar tests on the relationship of memory parameters across different frequencies have been introduced by Arteche (2002). The results in Theorem 2 and the Appendix suggest that these techniques remain valid for long memory signal plus noise models for memory parameters in the interval (0,0.5).

4 Finite samples and optimal bandwidth

The asymptotic properties of the Gaussian semiparametric estimator do not rely on the magnitude of the variance of ξ_t. For example the asymptotic variance does not depend on unknown parameters and is not affected by the magnitude of the variances of the signal and noise. However the finite sample performance will highly depend on the relation between these two variances; the spectral pole in v_t is harder to detect the larger σ_ξ^2 is relative to the variance of v_t. This effect of the added noise can be minimized by an adequate choice of the bandwidth m.

To analyse the finite sample performance of the Gaussian semiparametric estimate in a signal plus noise model I generate series of the form

$$y_t = v_t + \xi_t \tag{18}$$

for $(1-L)^d v_t = u_t$ and $\xi_t = \log \varepsilon_t^2$, for ε_t and u_t independent variables. I focus on standard long memory at zero frequency because this is the empirically more relevant and popular case. ε_t is standard normal in every case, corresponding to LMSV, and two different u_t are considered: $u_t = z_t$ and $\phi(L)u_t = z_t$ with $\phi(L) = 1 - 0.6L$ and z_t white noise normal with zero mean and variances $\sigma_z^2 = 1, 0.1$ and 0.05. I have also tried t_5 innovations which have bounded moments only to the fourth order, but the results do not differ from those with normal z_t and are not included. The smaller variances are closer to the values that have been empirically found when a LMSV model is fitted to many financial time series (see Jacquier et al., 1994, Breidt et al,. 1998 and Pérez and Ruiz, 2001, among others). For each of these processes three different d's are considered, 0.15, 0.30 and 0.45. I generate v_t using the Davies and Harte (1987) algorithm for the independent normal u_t and the truncated AR expansion of v_t in the other cases. Since LMSV models are usually applied to financial series which are generally quite large I just consider a sample size of 8192 observations and get the Gaussian semiparametric estimates of d using the observable data y_t. The estimates are obtained by a simple golden section search to the first derivative of $R(d)$ over $d \in [0.001, 0.499]$. For each situation I compare sample bias and mean square errors for 1000 replications. Following Henry and Robinson (1996) I consider an optimal bandwidth m_{opt} which minimizes an approximate mean square error (mse) of $\tilde{d}$ and compare it with the bandwidth m^* that minimizes a Monte Carlo mse. The bandwidth m_{opt} focuses on the local specification of the spectrum of v_t

$$f_v(\lambda) = C\lambda^{-2d}(1 + E_\beta \lambda^\beta + o(\lambda^\beta)).$$

The parametric long memory processes used in the literature comply $\beta \geq 1$. For example $\beta = 2$ in fractional ARIMA process and $\beta = 1$ for the parametric SCLM processes described in Arteche and Robinson (1999). In particular the generated v_t comply $\beta = 2$ and $C = \sigma_z^2/2\pi$ for u_t white noise and $C = \sigma_z^2/(2\pi\phi(1)^2)$ for the AR(1) u_t. Thus as $\lambda \to 0^+$

$$f_y(\lambda) = C\lambda^{-2d}\left(1 + \frac{\sigma_\xi^2}{2\pi C}\lambda^{2d} + O(\lambda^2)\right),$$

for $0 < d < 0.5$ and according to Henry and Robinson (1996) the bandwidth which minimizes an asymptotic approximation to the mse of the Gaussian semiparametric estimate is

$$m_{opt} = \left(\frac{(2d+1)^4}{16d^3 g^2 (2\pi)^{4d}}\right)^{\frac{1}{1+4d}} n^{\frac{4d}{1+4d}},$$

where $g = \sigma_\xi^2/2\pi C$. In this expression it is clear the influence of the variance of ξ_t on the performance of the estimator. The larger σ_ξ^2 with respect to C the more difficult is to detect the long range dependence of the series so that only frequencies very close to the origin contain a valuable information. The bandwidth m_{opt} is infeasible since depends on unknowns.

Estimation of nuisance parameters such as g is then required to get a plug-in version of m_{opt} as suggested by Henry and Robinson (1996). However the properties of these estimates are not known and incorrect estimates distort significantly the estimation of m_{opt}. Instead I found more adequate to minimize the objective function over a sensible set of bandwidths so that

$$\widehat{m} = \arg\min_m R(m, \widetilde{d}(m)).$$

Figure 1 shows sample bias and mse for the nine situations considered corresponding to $u_t = z_t$ normal white noise and different d and σ_z^2, for bandwidths $m = 1, 2, ..., 150$. The bias is large and negative and tends to increase with m, being larger the smaller σ_z^2 is. The mean square error also increases as σ_z^2 decreases and it gets minimal for a small number of frequencies, smaller the lower σ_z^2 is. A similar behaviour can be observed in figure 2 which shows bias and mse for the AR(1) u_t for bandwidths $m = 1, 2, ..., 600$, although in this case the positive bias caused by the AR(1) term tends to compensate the negative bias due to the added noise. Table 1 shows the bandwidth m^* that minimizes the Monte Carlo mse, m_{opt} (between round brackets) and the mean of the $\widehat{m}$ (between square brackets) for the different scenarios. Table 2 shows the mean square errors obtained with those bandwidths. Under white noise u_t, $\widehat{m}$ gives a smaller mean square error than the infeasible m_{opt} in six out of nine cases and in five of them when u_t is AR(1). The choice of the bandwidth $\widehat{m}$ tends to work worse for the higher memory parameter. It need also to be mentioned that, although $\widehat{m}$ has a good performance in mean, in many cases it gives a bandwidth quite far from the mean value, so that it is not a very reliable method to get an optimal m, and it is preferable to try different bandwidths.

I also implemented the variant of the log-periodogram regression of Geweke and Porter-Hudak (1983) as proposed by Robinson (1995a) and Arteche and Robinson (2000) and justified for local LMSV processes by Deo and Hurvich (2001). The Monte Carlo results (available upon request) for the series analysed in this section show higher mse´s in every situation and no lower bias than the Gaussian semiparametric estimate.

Figure 1: Bias and mse of Gaussian semiparametric estimates (u_t white noise normal).

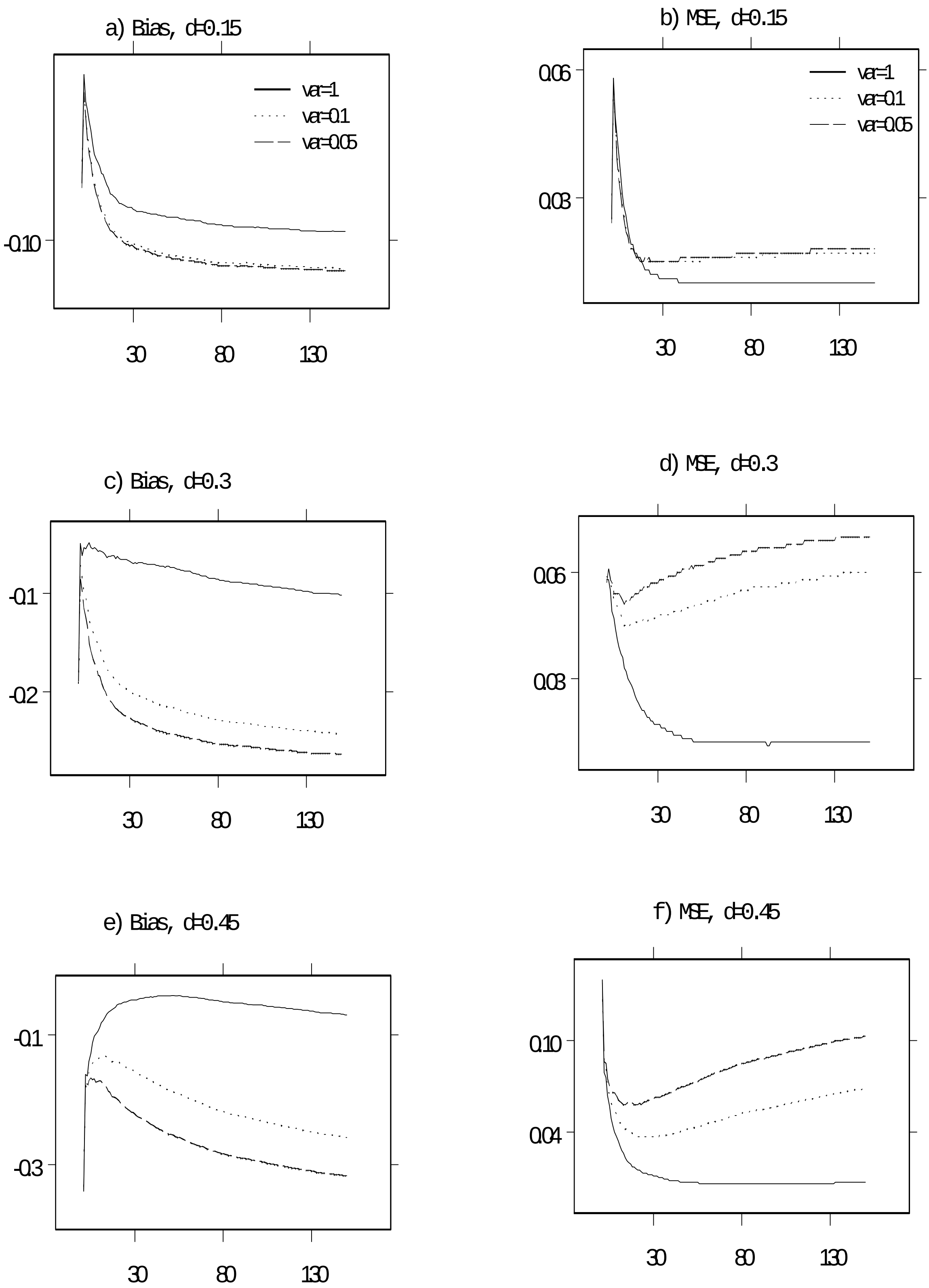


Figure 2: Bias and mse of Gaussian semiparametric estimates (u_t » AR(1))

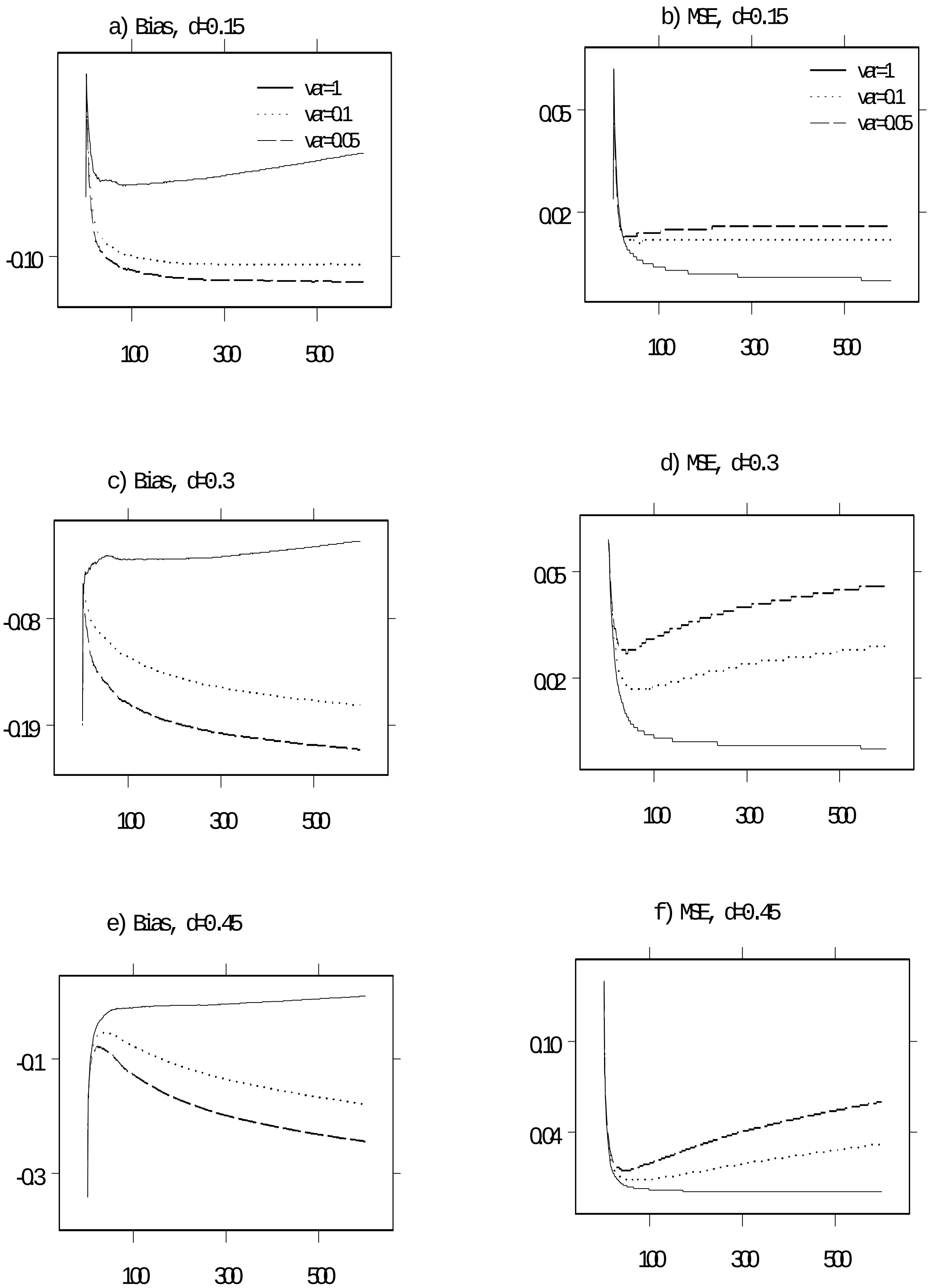


Table 1: Optimal bandwidths m^* (m_{opt}) [$\hat{m}$]

	$u_t \sim N(0,\sigma_z^2)$			$u_t \sim AR(1)$		
$\diagdown$ σ_z^2 / d $\diagdown$	**1**	**.1**	**.05**	**1**	**.1**	**.05**
.15	68 (24) [37]	30 (1) [13]	29 (1) [12]	599 (236) [434]	48 (13) [127]	33 (6) [56]
.3	93 (40) [118]	13 (5) [41]	11 (3) [25]	599 (213) [573]	55 (26) [501]	42 (14) [411]
.45	93 (70) [143]	23 (14) [123]	20 (8) [99]	582 (260) [594]	64 (50) [584]	45 (31) [565]

Table 2: MSE with optimal bandwidths m^* (m_{opt}) [$\hat{m}$]

	$u_t \sim N(0,\sigma_z^2)$			$u_t \sim AR(1)$		
$\diagdown$ σ_z^2 / d $\diagdown$	**1**	**.1**	**.05**	**1**	**.1**	**.05**
.15	.010 (.012) [.011]	.015 (.025) [.018]	.015 (.024) [.019]	.000 (.002) [.001]	.011 (.018) [.012]	.013 (.034) [.014]
.3	.011 (.014) [.012]	.044 (.052) [.049]	.051 (.058) [.056]	.000 (.002) [.000]	.017 (.021) [.028]	.027 (.034) [.043]
.45	.006 (.006) [.007]	.036 (.042) [.063]	.058 (.066) [.090]	.001 (.001) [.001]	.008 (.009) [.032]	.015 (.017) [.058]

5 Persistence in the volatility of IBEX35

This section analyses the persistence of the volatility of the Spanish stock index Ibex35 composed of the 35 more actively traded stocks. The frequency of the series is half-hourly and covers the period 1-10-93 to 22-3-96. The returns are constructed by first differencing the logarithm of the transaction prices of the last transaction every 30 minutes, omitting incomplete days. After this modification we get the series of intra-day returns x_t, $t = 1, ..., 7260$. The periodogram (in all the figures the periodogram is not evaluated at the origin so that global demeaning is not considered) of the series in Figure 3 shows no distinctive pattern which is in accord with the martingale difference hypothesis imposed by market efficiency. However the periodogram of $y_t = \log(x_t - \bar{x})^2$,where $\bar{x} = \sum x_t/7260$, in Figure 4 exhibits marked peaks at the origin and seasonal frequencies which reflect long run persistence and a strong seasonal behaviour. The strong seasonality can also be seen in the inverted J shape of the sample means of the 12 half-hourly intraday series, $s_a = \sum_{i=0}^{604} y_{a+12i}/605$ for $a = 1, ..., 12$, in Table 3.

Table 3: Intra-day sample means z_a

s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8	s_9	s_{10}	s_{11}	s_{12}
-2.15	-4.51	-5.21	-5.28	-5.59	-5.70	-5.68	-5.11	-5.17	-4.70	-4.07	-4.01

Since in the considered period the Spanish market opened at 11:00 and closed at 17:00 there are 12 observations per day so that the seasonal frequencies are $\omega_h = \pi h/6$, $h = 1, 2, ..., 6$. The peaks in the periodogram are exactly at those and no neighbouring frequencies which suggests deterministic seasonality in the form of seasonal dummies (Arteche, 2002). This first impression is corroborated by the periodogram of the deseasonalized -by regression on seasonal dummies- series in Figure 5 which shows no evidence of seasonality.

However, it is remarkable the power concentration around zero frequency, suggesting the possibility of standard long memory. Figure 6 shows Gaussian semiparametric estimates at zero frequency for $m = 6, ..., 600$. The higher values are obtained for smaller m. The low estimates obtained for higher bandwidths could be explained by the negative bias produced by an added noise as in the LMSV models.

6 Conclusion and extensions

The analysis of long memory in higher than first moments is gaining great importance especially in financial time series where often the raw series is not autocorrelated but there are some proxies of its volatility, such as squares, which show strong persistence. This paper adopts a local specification of the long memory in stochastic volatility model of Harvey (1998) and Breidt et al. (1998) and justify the estimation of the volatility memory parameter by means of the Gaussian semiparametric or local Whittle estimator. In fact, the consistency and asymptotic normality of the Gaussian semiparametric estimate is proved in a more general signal plus noise model. The conditions needed for consistency and asymptotic normality are milder than those required in Deo and Hurvich (2001) for the estimator based on the log-periodogram regression. For example Gaussianity is not necessary which is a convenient relaxation taking into account the characteristics of financial time series where these models

Figure 3: Periodogram of x_t (Ibex35 returns)

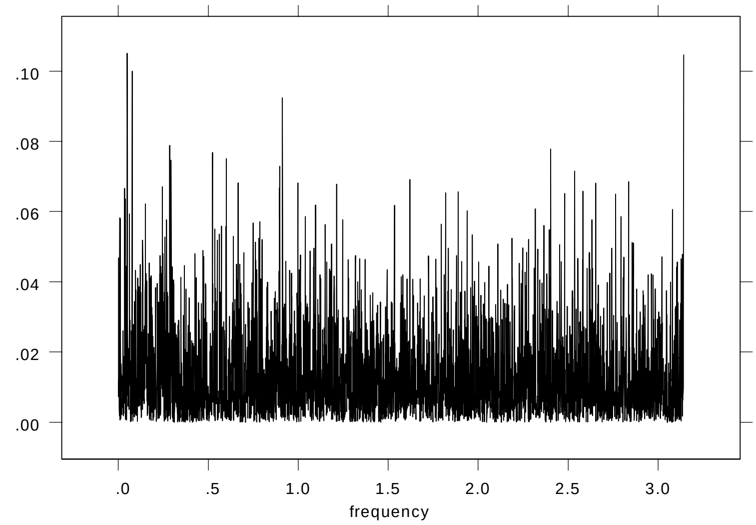


Figure 4: Periodogram of $y_t = \log(x_t - \overline{x})^2$

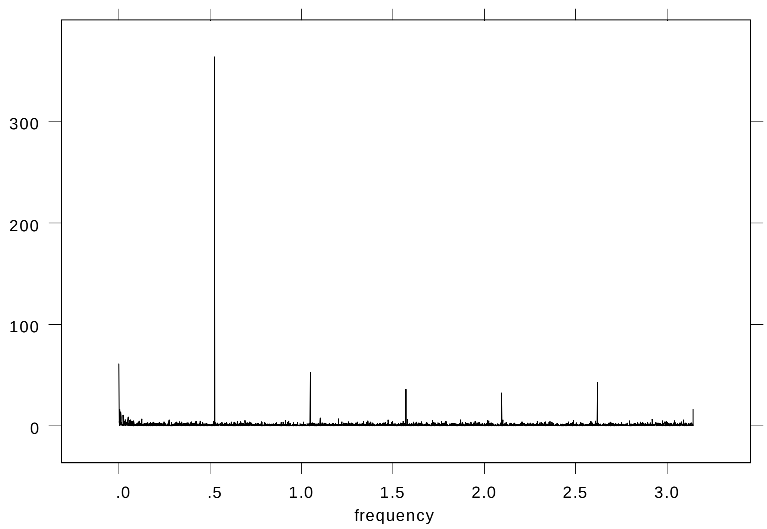


Figure 5: Periodogram of deseasonalized y_t:

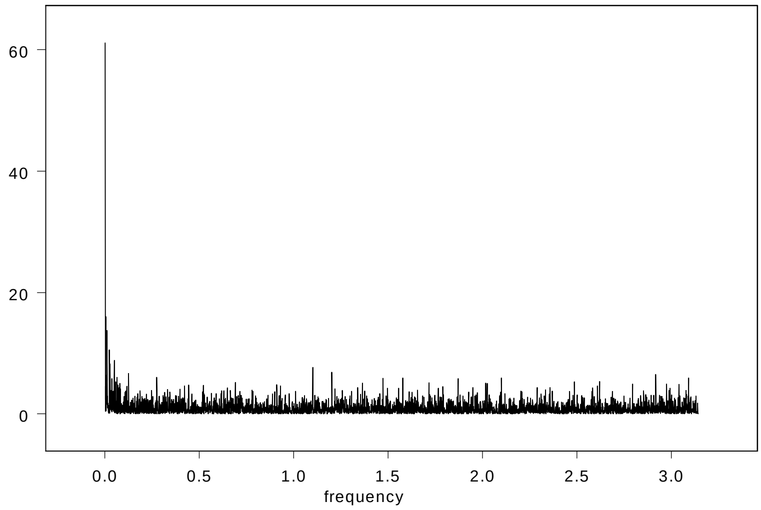
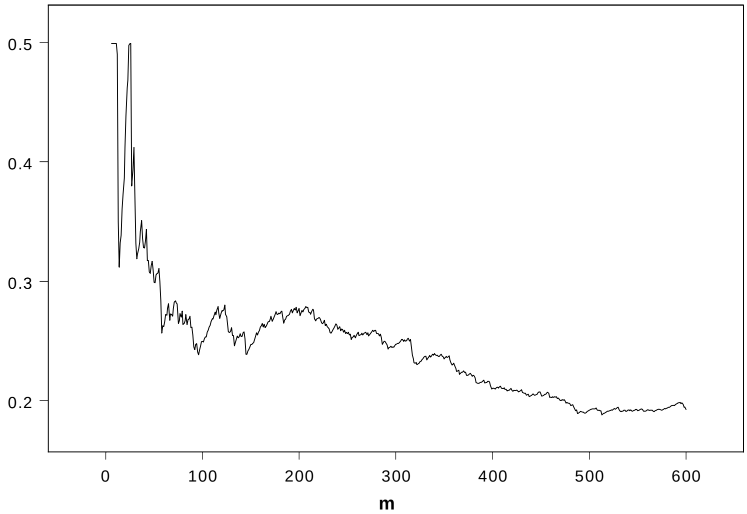


Figure 6: Gaussian semiparametric estimates at zero.



are more adequate. Moreover the Gaussian semiparametric estimate is more efficient asymptotically -with a lower asymptotic variance- and in finite samples -with a lower Monte Carlo mean square error-.

The added noise has a distorting effect on the estimates of the memory parameter of the signal. A suitable choice of the bandwidth is important to lessen its impact. A relevant topic for further research is the proposal of reliable data-driven approximations to an optimal bandwidth. Plug-in methods as those proposed by Henry and Robinson (1996) seem not adequate for two main reasons. First the infeasible m_{opt} is obtained by a heuristic asymptotic approximation to the true mse and it differs significantly from the bandwidth that minimizes a Monte Carlo mse. Secondly m_{opt} depends on d and nuisance parameters C and σ_ξ^2 and it is not clear how to obtain good estimates of these parameters. I have proposed a simple method based on a direct minimizing of the objective function. Although it works quite well in mean, extreme values are quite frequent and further research in this topic seems worthwhile. A related topic of interest is the reduction of the high bias of the estimates of the memory parameter in signal plus noise models.

Further research can also focus on multivariate extensions of LMSV models and its estimation and the analysis of common long-range components and cointegrating relationships between the volatility of different series.

Appendix: Technical Details

Assumptions for consistency:

A.1': For $\alpha \in (0, 2]$,

$$f_v(\omega \pm \lambda) = C\lambda^{-2d}(1 + O(\lambda^{\alpha})) \quad \text{as} \quad \lambda \to 0^+,$$

where $C \in (0, \infty)$, $0 < d < 0.5$.

A.2': In a neighbourhood $(-\delta, 0) \cup (0, \delta)$ of ω $f_v(\lambda)$ is differentiable and, as $\lambda \to 0^+$,

$$\left| \frac{\mathrm{d}}{\mathrm{d}\lambda} f_v(\omega \pm \lambda) \right| = O(\lambda^{-1-2d})$$

A.3': $v_t - Ev_1 = \sum_{j=0}^{\infty} \alpha_j \varepsilon_{t-j}$ and $\sum_{j=0}^{\infty} \alpha_j^2 < \infty$ where $E[\varepsilon_t | F_{t-1}] = 0$, $E[\varepsilon_t^2 | F_{t-1}] = 1$ for $t = 0, \pm 1, \pm 2, ...$, F_t is the σ-field generated by ε_s, $s \leq t$, and there exists a random variable ε such that $E\varepsilon^2 < \infty$ and for all $\eta > 0$ and some $\kappa < 1$, $P(|\varepsilon_t| > \eta) \leq \kappa P(|\varepsilon| > \eta)$.

A.4': As $n \to \infty$,

$$\frac{m}{n} + \frac{1}{m} \to 0.$$

Assumptions for asymptotic normality:

A.5': In a neighbourhood $(-\delta, 0) \cup (0, \delta)$ of ω, $\alpha(\lambda) = \sum_{k=0}^{\infty} \alpha_k e^{ik\lambda}$ is differentiable and

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} \alpha(\omega_i \pm \lambda) = O\left(\frac{|\alpha(\omega \pm \lambda)|}{\lambda} \right) \qquad \text{as } \lambda \to 0^+.$$

A.6': Assumption **A.3'** holds and

$$E(\varepsilon_t^3 | F_{t-1}) = \mu_3 \quad \text{and} \quad E(\varepsilon_t^4 | F_{t-1}) = \mu_4, \quad t = 0, \pm 1, ...,$$

for finite constants μ_3 and μ_4.

A.7': As $n \to \infty$

$$\frac{1}{m} + \frac{m^{1+2\alpha} (\log m)^2}{n^{2\alpha}} \to 0.$$

Proof of Theorem 1: The proof of the consistency is based on the proof of Theorem 1 in Arteche (2000) and Robinson (1995b) noting that

$$I_{yy}(\lambda) = I_{vv}(\lambda) + I_{\xi\xi}(\lambda) + I_{v\xi}(\lambda) + I_{\xi v}(\lambda) \tag{A.1}$$

where

$$I_{rp}(\lambda) = W_r(\lambda) \overline{W_p(\lambda)} = \frac{1}{2\pi n} \sum_{t=1}^{n} \sum_{s=1}^{n} r_t p_s \exp(i\lambda(s-t)), \qquad r, p = v, \xi,$$

are periodograms and cross-periodograms between v_t and ξ_t. Thus the only difference in the proof of the consistency with respect to Theorems 1 in Robinson (1995b) and Arteche (2000) comes out from the presence of $I_{\xi\xi}(\lambda) + I_{v\xi}(\lambda) + I_{\xi v}(\lambda)$ and we have only to prove that the terms involving $I_{\xi\xi}(\lambda)$, $I_{v\xi}(\lambda)$ and $I_{\xi v}(\lambda)$ are asymptotically negligible. Note that

$$\frac{I_{yy}(\omega + \lambda_j)}{g_j} - 1 = \frac{I_{vv}(\omega + \lambda_j)}{g_j} - 1 + \frac{I_{\xi\xi}(\omega + \lambda_j)}{g_j} + \frac{I_{v\xi}(\omega + \lambda_j)}{g_j} + \frac{I_{\xi v}(\omega + \lambda_j)}{g_j}$$

where $g_j = C|\lambda_j|^{-2d}$. By Theorem 10.3.2 in Brockwell and Davis (1991)

$$I_{\xi\xi}(\omega + \lambda_j) = O_p(1) \tag{A.2}$$

under assumption A.1, and the terms involving $I_{\xi\xi}(\omega+\lambda_j)/g_j$ are all $o_p(1)$. Since $|I_{v\xi}(\lambda)|^2 = I_{vv}(\lambda)I_{\xi\xi}(\lambda)$ we get from Theorem 5 in Arteche and Robinson (2000)

$$\frac{I_{v\xi}(\omega+\lambda_j)}{g_j} = O_p\left(\left[\frac{|j|}{n}\right]^d\right) \tag{A.3}$$

and the terms involving $I_{v\xi}(\omega+\lambda_j)/g_j$ and $I_{\xi v}(\omega+\lambda_j)/g_j$ are $O_p((m/n)^d) = o_p(1)$ for $d > 0$.The consistency thus follows directly from the corresponding theorems in Robinson (1995b) and Arteche (2000).

Proof of Theorem 2: The proof of the asymptotic normality follows from Theorem 2 in Robinson (1995b) and Theorem 2 in Arteche (2000). As $n \to \infty$,

$$p\lim \frac{\mathrm{d}^2 R(\bar{d})}{\mathrm{d}d^2} = 4A.4 \tag{19}$$

$$\sqrt{m\delta_\omega}\frac{\mathrm{d}R(d)}{\mathrm{d}d} = \frac{2}{\sqrt{\delta_\omega m}}\sum{}' q_j \frac{I_{yy}(\omega+\lambda_j)}{g_j}(1+o_p(1))A.5 \tag{20}$$

$$= \frac{2}{\sqrt{\delta_\omega m}}\sum{}' q_j \frac{I_{vv}(\omega+\lambda_j)}{g_j}(1+o_p(1))A.6 \tag{21}$$

$$\xrightarrow{d} N(0,4)A.7 \tag{22}$$

where $|\bar{d}-d| \le |\tilde{d}-d|$ and $q_j = \log|j| - (\delta_\omega m)^{-1}\sum{}' \log|j|$. The result in $(A.4)$ comes from similar operations to those in the proof of the consistency. The equality in $(A.5)$ and the convergence in $(A.7)$ come directly from Robinson (1995b), formula (4.11), and formulae (3.9) and (3.10) in Arteche (2000). Finally the equality in $(A.6)$ comes from $(A.1)$ and $(A.2)$ so that

$$\frac{1}{\sqrt{\delta_\omega m}}\sum{}' q_j \frac{I_{\xi\xi}(\omega+\lambda_j)}{g_j} = O_p\left(\log m \frac{m^{\frac{1}{2}+2d}}{n^{2d}}\right) = o_p(1) \tag{A.8}$$

under assumption A.2, and the fact that

$$\frac{1}{\sqrt{\delta_\omega m}}\sum{}' q_j \frac{I_{v\xi}(\omega+\lambda_j)}{g_j} = o_p(1). \tag{A.9}$$

To prove $(A.9)$ it is sufficient to show that

$$E\left[\frac{1}{\sqrt{\delta_\omega m}}\sum{}' q_j \frac{I_{v\xi}(\omega+\lambda_j)}{g_j}\right]^2 = o(1) \tag{A.10}$$

which implies $(A.9)$ since $EI_{v\xi}(\lambda) = 0$ for all λ. The left hand side of $(A.10)$ is

$$\frac{1}{\delta_\omega m} \sum_j{}' \frac{q_j}{g_j} \sum_k{}' \frac{q_k}{g_k} E(W_{vj}\overline{W}_{\xi j} W_{\xi k}\overline{W}_{vk}) \tag{A.11}$$

where

$$W_{vj} = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} v_t \exp(-it(\omega + \lambda_j))$$

$\overline{W}_{vj}$ is its complex conjugate and the rest of discrete Fourier transforms are defined similarly. By independence of v and $\xi (A.11)$ is equal to

$$\frac{1}{\delta_\omega m} \sum_j{}' \frac{q_j^2}{g_j} E\left(\frac{I_{vv}(\omega + \lambda_j)}{g_j}\right) EI_{\xi\xi}(\omega + \lambda_j) A.12 \tag{23}$$

$$+\frac{1}{\delta_\omega m} \sum_j{}' \sum_{k \neq j}{}' \frac{q_j}{g_j}\frac{q_k}{g_k} E(W_{vj}\overline{W}_{vk}) E(\overline{W}_{\xi j} W_{\xi k}) A.13 \tag{24}$$

Now $(A.12)$ is

$$\begin{aligned} & O\left(\left[\frac{m}{n}\right]^{2d} \frac{1}{\delta_\omega m} \sum_j{}' q_j^2 \left[1 + \frac{\log|j|}{|j|} + \left(\frac{|j|}{n}\right)^{\alpha}\right]\right) \\ = \; & O\left(\left[\frac{m}{n}\right]^{2d} \log m\right) = o(1) \end{aligned}$$

under assumption A.2, because of $(A.2)$, Theorem 5 in Arteche and Robinson (2000) and

$$\frac{1}{\delta_\omega m} \sum_j{}' q_j^2 = 1 + O\left(\frac{\log^2 m}{m}\right).$$

Regarding $(A.13)$, we have

$$E(\overline{W}_{\xi j} W_{\xi k}) = \frac{\sigma_\xi^2}{2\pi n} \sum_{t=1}^{n} \exp(it(\lambda_j - \lambda_k)) = 0 \qquad \text{for } k \neq j$$

if ξ_t is white noise and, $E(\overline{W}_{\xi j} W_{\xi k}) = o(1)$ for $j \neq k$ under assumption A.1 so that $(A.13)$ is

$$o_p\left(\frac{1}{m} \sum_j{}' \sum_{k<j}{}' \frac{q_j}{\sqrt{g_j}} \frac{q_k}{\sqrt{g_k}} \frac{\log|j|}{|k|}\right) = o_p\left(\frac{(\log m)^3}{mn^{2d}} \sum_j{}' |j|^d \sum_{k<j}{}' |k|^{d-1}\right) = o_p\left((\log m)^3 \left[\frac{m}{n}\right]^{2d}\right)$$

which is $o_p(1)$ under assumption A.2 and because of Theorem 5 in Arteche and Robinson (2000). Thus $(A.9)$ holds and the proof is completed.

If ξ_t were LM with spectral density function satisfying

$$f_\xi(\omega \pm \lambda) = C_\zeta \lambda^{-2c}(1 + O(\lambda^\beta)) \quad \text{and} \quad \left|\frac{\mathrm{d}}{\mathrm{d}\lambda} f_\xi(\omega \pm \lambda)\right| = O(\lambda^{-1-2c}) \qquad \text{as } \lambda \to 0^+$$

for $\beta \in (0,2]$ and $0 < c < d$ we get from Theorem 5 in Arteche and Robison that

$$E\left|\frac{I_{\xi\xi}(\omega+\lambda_j)}{g_j}\right| \leq cte\left(\frac{|j|}{n}\right)^{2(d-c)}$$

and the right hand side of $(A.8)$ is in this case

$$O_p\left(\log m \frac{m^{\frac{1}{2}+2(d-c)}}{n^{2(d-c)}}\right).$$

Similarly we get by Theorem 5 in Arteche and Robinson (2000)

$$\begin{aligned}
\frac{1}{\delta_\omega m}\sum_j{}' \frac{q_j^2}{g_j} E\left(\frac{I_{vv}(\omega+\lambda_j)}{g_j}\right) EI_{\xi\xi}(\omega+\lambda_j) &= O\left(\left[\frac{m}{n}\right]^{2(d-c)} \log^2 m\right) \\
\frac{1}{\delta_\omega m}\sum_j{}' \sum_{k\neq j}{}' \frac{q_j}{g_j}\frac{q_k}{g_k} E(W_{vj}\overline{W}_{vk})E(\overline{W}_{\xi j}W_{\xi k}) &= O\left(\left[\frac{m}{n}\right]^{2(d-c)} \frac{\log^4 m}{m}\right).
\end{aligned}$$

References

[1] Andersen, T.G. and T. Bollerslev, 1997. Intraday periodicity and volatility persistence in …nancial markets. Journal of Empirical Finance 4, 115-158.

[2] Arteche, J., 1998. Seasonal and Cyclical Long Memory in Time Series. Ph.D. thesis, University of London.

[3] Arteche, J., 2000a. Gaussian semiparametric estimation in seasonal or cyclical long memory time series. Kybernetika 36, 279-310.

[4] Arteche, J., 2000b. Semiparametric robust tests on seasonal or cyclical long memory time series. Preprint.

[5] Arteche, J. and P.M. Robinson, 1999. Seasonal and cyclical long memory, in S. Ghosh ed., Asymptotics, Nonparametrics and Time Series, (Marcel Dekker Inc., New York) 115-148.

[6] Arteche, J. and P.M. Robinson, 2000. Semiparametric inference in seasonal and cyclical long memory processes. Journal of Time Series Analysis 21, 1-25.

[7] Baillie, R.T., T. Bollerslev and H.O. Mikkelsen, 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 74, 3-30.

[8] Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307-327.

[9] Bollerslev, T. and H.O. Mikkelsen, 1996. Modeling and pricing long memory in stock market volatility. Journal of Econometrics 73, 151-184.

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

[11] Carlin, J.B. and A.P. Dempster, 1989. Sensitivity analysis of seasonal adjustments: empirical case studies. Journal of the American Statistical Association 84, 6-20.

[12] Chan, N.H. and N. Terrin, 1995. Inference for unstable long-memory processes with applications to fractional unit root autoregressions. The Annals of Statistics 23, 1662-1683.

[13] Chan, N.H. and C.Z. Wei, 1988. Limiting distributions of least squares estimates of unstable autoregressive processes. The Annals of Statistics 16, 367-401.

[14] Cheung, Y.W. and F.X. Diebold, 1994. On maximum likelihood estimation of the differencing parameter of fractionally integrated noise with unknown mean. Journal of Econometrics 62, 301-316.

[15] Chung, C.F., 1996. Estimating a generalized long memory process. Journal of Econometrics 73, 237-259.

[16] Davies, R.B. and D.S. Harte (1987). Tests for Hurst e¤ect. Biometrika 74, 95-101.

[17] Dahlahaus, R., 1989. E¢cient parameter estimation for self-similar processes. The Annals of Statistics 17, 1749-1766.

[18] Deo, R.S. and C.M. Hurvich, 1999. On the log periodogram regression estimator of the memory parameter in long memory stochastic volatility models. Preprint.

[19] Ding, Z. and C.W.J. Granger, 1996. Modeling volatility persistence of speculative returns: A new approach. Journal of Econometrics 73, 185-215.

[20] Ding, Z., C.W.J. Granger, and R.F. Engle, 1993. A long memory property of stock market returns and a new model. Journal of Empirical Finance 1, 83-106.

[21] Engle, R.F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom in‡ation. Econometrica 50, 987-1007.

[22] Engle, R.F. and T. Bollerslev, 1986. Modelling the persistence of conditional variances. Econometric Reviews 5, 1-50.

[23] Fox, R. and M.S. Taqqu, 1986. Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. The Annals of Statistics 14, 517-532.

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

[25] Giraitis, L. and R. Leipus, 1995. A generalized fractionally di¤erencing approach in long-memory modelling. Lietuvos Matematikos Rinkinys 35, 65-81.

[26] Giraitis, L., P.M. Robinson and A. Samarov, 1997. Rate optimal semiparametric estimation of the memory parameter of the Gaussian time series with long range dependence. Journal of Time Series Analysis 18, 49-60.

[27] Giraitis, L. and D. Surgailis, 1990. A central limit theorem for quadratic forms in strongly dependent linear variables and its application to asymptotic normality of Whittle's estimate. Probability Theory and Related Fields 86, 87-104.

[28] Granger, C.W.J. and R. Joyeux, 1980. An introduction to long-memory time series models and fractional di¤erencing. Journal of Time Series Analysis 4, 221-238.

[29] Gray, H.L., N.F. Zhang and W.A. Woodward, 1989. On generalized fractional processes. Journal of Time Series Analysis 10, 233-257.

[30] Gray, H.L., N.F. Zhang and W.A. Woodward, 1994. On generalized fractional processes-a correction. Journal of Time Series Analysis 15, 561-562.

[31] Harvey, A.C., 1993. Long memory in stochastic volatitility. Manuscript, London School of Economics.

[32] Harvey, A.C., E. Ruiz and N. Shephard, 1994. Multivariate stochastic variance models. Review of Economic Studies 61, 247-264.

[33] Henry, M. and P.M. Robinson, 1996. Bandwidth choice in Gaussian semiparametric estimation of long range dependence. Lecture Notes in Statistics 115 ,(Springer-Verlag, New York).

[34] Heyde, C.C. and R. Gay, 1993. Smoothed periodogram asymptotics and estimation for processes and …elds with possible long-range dependence. Stochastic Processes and their Applications 45, 169-182.

[35] Hosking, J.R.M., 1981. Fractional di¤erencing. Biometrika 68, 165-176.

[36] Jacquier, E., N.G. Polson and P.E. Rossi, 1994. Bayesian analysis of stochastic volatility models. Journal of Business and Economic Statisitics 12, 371-417.

[37] Jonas, A.J., 1983. Persistent Memory Random Processes. Ph.D. Thesis, Harvard Univeristy.

[38] Kunsch, H., 1987. Statistical aspects of self-similar processes, in Y. Prohorov and V.V. Sazanov eds., Proceedings of the First World Congress of the Bernoulli Society, (VNU Science Press, Utrecht) 67-74.

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

[40] Nelson, D.B., 1990. Stationarity and persistence in the GARCH(1,1) model. Econometric Theory 6, 318-334.

[41] Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59, 347-370.

[42] Pérez, A. and E. Ruiz. 2001. Finite sample properties of a QML estimator of stochastic volatility models with long memory. Economic Letters 70, 157-164.

[43] Robinson, P.M., 1991. Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. Journal of Econometrics 47, 67-84.

[44] Robinson, P.M., 1994. E¢cient tests of non-stationary hypotheses. Journal of the American Statistical Association 89, 1420-1437.

[45] Robinson, P.M., 1995a. Log-periodogram regression of time series with long-range dependence. The Annals of Statistics 23, 1048-1072.

[46] Robinson, P.M., 1995b. Gaussian semiparametric estimation of long-range dependence. The Annals of Statistics 23, 1630-1661.

[47] Robinson, P.M. and P. Za¤aroni, 1997. Nonlinear time series with long memory: a model for stochastic volatility. LSE STICERD working paper EM/97/320.

[48] Taylor, S.J., 1986. Modelling Financial Time Series. (John Wiley and Sons, Chichester).

[49] Whittle, P. 1953. Estimation and information in stationary time series. Arkiv för Matematik 2, 423-434.