Augmented log-periodogram regression in long memory signal plus noise models

J. Arteche*

Departamento de Econometría y Estadística
University of the Basque Country (UPV-EHU)
Bilbao 48015, Spain
Tl: (34) 94 601 3852
Fax: (34) 94 601 3754
Email: etpargoj@bs.ehu.es

6th November 2003

*Research supported by the University of the Basque Country grant 9/UPV 00038.321-13503/2001 and the Spanish *Ministerio de Ciencia y Tecnología* and *FEDER* grant BEC2003-02028.

Running head: Augmented log periodogram regression

Corresponding address:

Josu Arteche
Departamento de Econometría y Estadística
Facultad de Ciencias Económicas y Empresariales
University of the Basque Country (UPV-EHU)
Bilbao 48015
Spain
Tl: (34) 94 601 3852
Fax: (34) 94 601 3754
Email: etpargoj@bs.ehu.es

Abstract

The estimation of the memory parameter in perturbed long memory series has recently attracted attention motivated especially by the strong persistence of the volatility of many financial and economic time series and the use of Long Memory in Stochastic Volatility (LMSV) processes to model such a behaviour. This paper proposes an extension of the log periodogram regression which explicitly accounts for the added noise. Contrary the the non linear log periodogram regression of Sun and Phillips (2003), no linear approximation of the logarithmic term which accounts for the noise is used. This produces a reduction of the bias and increases the asymptotic efficiency in long memory signal plus noise series. Asymptotic and finite sample properties of the estimator are analyzed. Finally an application to the Spanish stock index Ibex35 is included.

1 INTRODUCTION

The estimation of the memory parameter in perturbed long memory processes has become a subject of increasing interest motivated especially by the strong persistence found in the volatility of many financial and economic series. Alternatively to the different extensions of ARCH and GARCH models, the Long Memory in Stochastic Volatility (LMSV) has proved a useful tool to model such a strong persistent volatility. A logarithmic transformation of the squared series becomes a long memory process perturbed by an additive noise where the long memory signal corresponds to the volatility of the original series. As a result estimation of the memory parameter of the volatility component in LMSV corresponds to a problem of estimation in a long memory signal plus noise model. Several estimation techniques have been proposed in this context (Harvey(1998), Breidt et al.(1998), Deo and Hurvich (2001), Arteche (2003), Sun and Phillips (2003)).

The perturbed long memory series recently considered in the literature are of the form,

$$z_t = \mu + y_t + u_t \tag{1}$$

where μ is a finite constant, y_t is a long memory (LM) process such that its spectral density satisfies

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

for a positive finite constant C, $\alpha \in [1, 2]$ and $0 < d < 0.5$, and u_t is a weakly dependent process. The LMSV model considers u_t a non normal white noise but in a more general signal plus noise model u_t can be a serially dependent process as in Arteche (2003) and Sun and Phillips (2003). The constant α determines the smoothness of the spectral density of y_t around the origin. In the standard fractional ARIMA processes $\alpha = 2$. We also consider the cases $2 > \alpha \geq 1$ which correspond for example to the seasonal or cyclical long memory processes of Arteche and Robinson (1999) where $\alpha = 1$. The condition of positive memory $0 < d < 0.5$ is usually imposed when dealing with frequency domain estimation in perturbed long memory processes and it guarantees the asymptotic equivalence between spectral densities of y_t and z_t. For u_t uncorrelated with y_t the spectral density of z_t is

$$f_z(\lambda) = f_y(\lambda) + f_u(\lambda) = C\lambda^{-2d}(1 + O(\lambda^{\alpha})) + f_u(\lambda) \sim C\lambda^{-2d} \qquad \text{as } \lambda \to 0 \tag{3}$$

and z_t inherits the memory properties of y_t. The semiparametric estimators of d based on the partial specification of the spectral density in (2) can also be applied in perturbed LM series as z_t. Two semiparametric estimators are mostly used: the log periodogram regression of Geweke and Porter-Hudak (1983) and the variant of Robinson (1995a), and the local Whittle or Gaussian semiparametric estimator of Robinson (1995b). Based on (2) Deo and Hurvich (2001) and Arteche (2003) proved the validity of both techniques in certain perturbed LM processes.

The version of Robinson (1995a) of the log periodogram regression estimator (LPE) in a fully observable LM series is based on the least squares regression

$$\log I_{yj} = a + d(-2\log\lambda_j) + v_j, \qquad j = 1, ..., m,$$

where

$$I_{yj} = I_y(\lambda_j) = \frac{1}{2\pi n}\left|\sum_{t=1}^{n} y_t \exp(-i\lambda_j t)\right|^2$$

is the periodogram of the series y_t, $t = 1, ..., n$, at Fourier frequency $\lambda_j = 2\pi j/n$ and m is the bandwidth such that at least $m^{-1} + mn^{-1} \to 0$ as $n \to \infty$. The original regressor proposed by Geweke and Porter-Hudak was $-2\log(2\sin\frac{\lambda_j}{2})$ instead of $-2\log\lambda_j$ but both are asymptotically equivalent and the differences between using one or another are minimal. The motivation of this estimator is the log linearization in (2) such that

$$\log I_{yj} = a + d(-2\log\lambda_j) + U_{yj} + O(\lambda_j^{\alpha}), \quad j = 1, 2, ..., m, \tag{4}$$

where $a = \log C - c_0$, $c_0 = 0.577216...$ is Euler's constant and $U_{yj} = \log(I_{yj} f_y^{-1}(\lambda_j)) + c_0$. The bias of the least squares estimate of d is dominated by the $O(\lambda_j^{\alpha})$ term which is not explicitly considered in the regression such that a bias of order $O(\lambda_m^{\alpha})$ arises (Hurvich et al. (1998))

The main rival semiparametric estimator of the LPE is the local Whittle or Gaussian semiparametric estimator (GSE) of Robinson (1995b) defined as the minimizer of

$$R(d) = \log\tilde{C}(d) - \frac{2d}{m}\sum_{j=1}^{m}\log\lambda_j, \quad \tilde{C}(d) = \frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2d} I_{yj} \tag{5}$$

over a compact set. This estimator has the computational disadvantage of requiring nonlinear optimization but it is more efficient than the log periodogram regression. However both share important affinities as described in Robinson and Henry (2003). In particular the bias is in both cases of order $O(\lambda_m^\alpha)$ (see Henry and Robinson (1996) for the GSE).

Both estimators preserve the consistency and asymptotic normality when applied to perturbed long memory series (Deo and Hurvich (2001) and Arteche (2003)). In this case

$$f_z(\lambda) = C\lambda^{-2d}\left(1 + \frac{f_u(\lambda)}{C}\lambda^{2d} + O(\lambda^\alpha)\right) \qquad \text{as } \lambda \to 0 \tag{6}$$

with $f_u(\lambda)$ positive and bounded. The leading term of the bias is in both estimators $f_u(\lambda)C^{-1}\lambda^{2d}$ which is the dominant part not considered explicitly in the estimation. Then the bias is of order $O(\lambda_m^{2d})$ which can be quite severe, especially if d is low. Correspondingly the asymptotic normality requires at least $m^{1+4d}n^{-4d} \to 0$ as $n \to \infty$ which limits the size of the bandwidth and consequently the asymptotic efficiency of both estimators.

In order to reduce the bias of the GSE Hurvich et al. (2003) suggested to incorporate explicitly a $\beta\lambda_j^{2d}$ term in the estimation to take into account the effect of the added noise on the spectral density of z_t and proposed a modified Gaussian semiparametric estimator (MGSE) defined as

$$(\tilde{d}_M, \tilde{\beta}_M) = \arg\min_{\Delta\times\Theta} R(d,\beta) \tag{7}$$

where $\Theta = [0, \Theta_1]$, $\Theta_1 < \infty$, $\Delta = [\Delta_1, \Delta_2]$, $0 < \Delta_1 < \Delta_2 < 1/2$,

$$R(d,\beta) = \log\left(\frac{1}{m}\sum_{j=1}^{m}\frac{\lambda_j^{2d}I_{yj}}{1+\beta\lambda_j^{2d}}\right) + \frac{1}{m}\sum_{j=1}^{m}\log\{\lambda_j^{-2d}(1+\beta\lambda_j^{2d})\}$$

In Hurvich et al. (2003) u_t is $iid(0,\sigma_u^2)$ so that $f_u(\lambda) = \sigma_u^2(2\pi)^{-1}$ and $\beta = \sigma_u^2(2\pi C)^{-1}$. Including this term in the estimation procedure the upper bound in the bandwidth is relaxed to comply

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

as $n \to \infty$ which allows a gain in asymptotic efficiency. In the fractional ARIMA processes the MGSE achieves a rate of convergence arbitrarily close to $n^{2/5}$ which is the upper bound of the rate of convergence of the Gaussian semiparametric estimator in the absence of additive noise. However with an additive noise the best possible rate of convergence achieved by the GSE is $n^{2d/(4d+1)}$. Regarding the bias, the MGSE has a bias of order $O(\lambda_m^\alpha)$ instead of $O(\lambda_m^{2d})$ which is the bias of the GSE in the presence of an additive noise.

Sun and Phillips (2003) extended the log periodogram regression in a similar manner. From (6) with $f'_u(0) = 0$

$$
\begin{aligned}
\log I_{zj} &= \log C - c_0 + d(-2\log\lambda_j) + \log\left(1 + \frac{f_u(\lambda)}{C}\lambda_j^{2d} + O(\lambda_j^{\alpha})\right) + U_{zj} \\
&= \log C - c_0 + d(-2\log\lambda_j) + \log\left(1 + \frac{f_u(0)}{C}\lambda_j^{2d}\right) + \log\left(1 + \frac{O(\lambda_j^{\alpha})}{1 + \frac{f_u(0)}{C}\lambda_j^{2d}}\right) + U_{zj} \\
&= \log C - c_0 + d(-2\log\lambda_j) + \log\left(1 + \frac{f_u(0)}{C}\lambda_j^{2d}\right) + O(\lambda_j^{\alpha}) + U_{zj} \qquad (8) \\
&= \log C - c_0 + d(-2\log\lambda_j) + \frac{f_u(0)}{C}\lambda_j^{2d} + O(\lambda_j^{\alpha^*}) + U_{zj}
\end{aligned}
$$

where $\alpha^* = \min(4d, \alpha)$. Sun and Phillips (2003) proposed a non linear log periodogram regression

$$\log I_{zj} = a + d(-2\log\lambda_j) + \beta\lambda_j^{2d} + U_{zj} \qquad (9)$$

for $\beta = f_u(0)/C$, such that the non linear log periodogram regression estimator (NLPE) is defined as

$$(\hat{a}, \hat{d}, \hat{\beta}) = \arg\min \sum_{j=1}^{m} (\log I_{zj} - a + d(2\log\lambda_j) - \beta\lambda_j^{2d})^2 \qquad (10)$$

The bias of $\hat{d}$ is of order $O(\lambda_m^{\alpha^*})$ which is produced by the $O(\lambda_j^{\alpha^*})$ omitted in the regression in (9). Correspondingly the upper bound of m for the asymptotic normality is $O(n^{2\alpha^*/(2\alpha^*+1)})$. Sun and Phillips (2003) only consider the case $\alpha = 2$ so that $\alpha^* = 4d$ and the behaviour of m is restricted to be $O(n^{8d/(8d+1)})$ with a bias of order $O(\lambda_m^{4d})$, but the extension to $\alpha < 2$ is straightforward. The asymptotic efficiency of the NLPE is higher than in the standard LPE but lower than the asymptotic efficiency of the MGSE when $\alpha > 4d$. The reason of this behaviour is the approximation of the log expression in (8). This approach has been used by Andrews and Geggenberger (2003) in their bias reduced log periodogram regression in order to get a linear regression model. However, the regression model of Sun and Phillips (2003) is still non linear and the approximate linearization of the logarithmic term does not imply any computational advantage. Instead, noting (8) we propose the following non linear regression model

$$\log I_{zj} = a + d(-2\log\lambda_j) + \log(1 + \beta\lambda_j^{2d}) + U_{zj} \qquad (11)$$

which only leaves an $O(\lambda_j^{\alpha})$ term out of explicit consideration.

Section 2 considers the bias of the periodogram as an approximation of the spectral density of the LM signal. This bias is directly related to the bias of LPE and GSE of the memory parameter in a perturbed LM process. Section 3 analyzes the asymptotic properties of the estimator based in the regression model in equation (11) that we call the augmented log periodogram estimator (ALPE) . Section 4 considers its finite sample behaviour and compares it with the LPE and NLPE in LMSV models. Section 5 shows an application to a Ibex35 stock index series. Finally section 6 concludes. Technical details are placed in the Appendix.

2 ASYMPTOTIC BIAS OF THE PERIODOGRAM

The properties of the different estimators of d depend on the adequacy of the approximation of the periodogram to the spectral density. Hurvich and Beltrao (1993), Robinson (1995a) and Arteche and Velasco (2003) in an asymmetric long memory context, observed that the asymptotic relative bias of the periodogram as an approximation of the spectral density produces the bias typically encountered in semiparametric estimates of the memory parameters.

Deo and Hurvich (2001), Crato and Ray (2002) and Arteche (2003) found that the bias is quite severe in perturbed long memory series if the added noise is not explicitly considered in the estimation. This bias is caused by the poor approximation of the periodogram of z_t to the spectral density of the signal y_t. This section analyzes the asymptotic behaviour of the periodogram of the observable series z_t as an approximation of the spectral density of the latent signal y_t when evaluated at the Fourier frequency λ_j for both j fixed and increasing with n as $n \to \infty$.

Consider the following assumptions:

A.1: z_t in (1) is a long memory signal plus noise process with y_t an LM process with spectral density function in (2) with $d < 0.5$ and u_t is stationary with positive and bounded continuous spectral density function $f_u(\lambda)$.

A.2: y_t and u_t are independent.

Theorem 1 *Let z_t satisfy assumptions A.1 and A.2 and define*

$$L_j(d) = E\left[\frac{I_{zj}}{C\lambda_j^{-2d}}\right].$$

Then, considering j fixed:

$$L_j(d) = A_{1j} + A_{2j} + o(n^{-2d})$$

where

$$\lim_{n\to\infty} A_{1j} = \int_{-\infty}^{\infty} \psi_j(\lambda)\left|\frac{\lambda}{2\pi j}\right|^{-2d} d\lambda$$

and

$$\lim_{n\to\infty} n^{2d} A_{2j} = \int_{-\infty}^{\infty} \psi_j(\lambda)\frac{f_u(0)}{C(2\pi j)^{-2d}} d\lambda$$

where

$$\psi_j(\lambda) = \frac{2}{\pi}\frac{\sin^2\frac{\lambda}{2}}{(2\pi j-\lambda)^2}.$$

Remark 1: In the LMSV case $f_u(0) = \sigma_\xi^2/2\pi$. The influence of the noise is clear here, the larger the variance of the noise the higher the relative bias of the periodogram. This explains the high bias of semiparametric estimates in LMSV models under a low signal to noise ratio in Crato and Ray (2002) and Arteche (2003).

Remark 2: When $d < 0$ the bias increases without limit as n increases. This justifies the difficulties encountered when estimating a negative d in perturbed long memory series (Deo and Hurvich (2001) and Arteche (2003)).

It is also interesting to consider the asymptotic bias of the periodogram at Fourier frequencies with j increasing with n.

Theorem 2 *Let z_t satisfy assumptions A.1 and A.2, and consider a sequence of positive integers $j = j(n)$ such that $j/n \to 0$ as $n \to \infty$. Then*

$$L_j(d) = 1 + O\left(\frac{\log j}{j} + \lambda_j^{\min(\alpha,2d)}\right)$$

Proof: The only variation with the proof of Theorem 2 in Robinson (1995a) comes from the difference between $f_z(\lambda_j)$ and $C\lambda_j^{-2d}$ which by assumptions A.1 and A.2 is

$$f_z(\lambda_j) - C\lambda_j^{-2d} = f_y(\lambda_j) + f_u(\lambda_j) - C\lambda_j^{-2d}$$

and by assumption A.1,

$$\frac{f_z(\lambda_j)}{C\lambda_j^{-2d}} = 1 + O\left(\lambda_j^{\min(\alpha,2d)}\right). \quad \square$$

The second term in the right hand side of the equality in Theorem 2 can be guaranteed to be negligible only under positive memory. The spectral density of z_t only inherits the behavior of the spectrum of y_t if $d > 0$. This makes difficult the estimation of the memory parameter of the signal for $d < 0$. In fact, semiparametric estimates of d in perturbed LM series have been only proposed for positive memory avoiding the negative d case.

3 AUGMENTED LOG PERIODOGRAM REGRESSION

The augmented log periodogram estimator (ALPE) is defined as

$$(\tilde{a}, \tilde{d}, \tilde{\beta}) = \arg\min \sum_{j=1}^{m} (\log I_{zj} - a + d(2\log\lambda_j) - \log(1 + \beta\lambda_j^{2d}))^2 \tag{12}$$

under the constraint $\beta \geq 0$. Concentrating the constant a out

$$(\tilde{d}, \tilde{\beta}) = \arg\min_{\Delta\times\Theta} Q(d, \beta) \tag{13}$$

where

$$Q(d, b) = \sum_{j=1}^{m} (\log I_{zj}^* + d(2\log\lambda_j)^* - \log^*(1 + \beta\lambda_j^{2d}))^2$$

and for a general ξ_t we use the notation $\xi_t^* = \xi_t - \bar{\xi}$ where $\bar{\xi} = \sum \xi_t / n$.

The first order conditions of this minimization problem are

$$\begin{aligned} S(\tilde{d}, \tilde{\beta}) &= (0, \Lambda)' \\ \Lambda\beta = 0 & \end{aligned}$$

where

$$S(d, \beta) = \sum_{j=1}^{m} \begin{pmatrix} x_{1j}^*(d, \beta) \\ x_{2j}^*(d, \beta) \end{pmatrix} W_j(d\beta)$$

whith

$$\begin{aligned} x_{1j}(d, \beta) &= 2\left(1 - \frac{\beta\lambda_j^{2d}}{1 + \beta\lambda_j^{2d}}\right)\log\lambda_j\ , \\ x_{2j}(d, \beta) &= -\frac{\lambda_j^{2d}}{1 + \beta\lambda_j^{2d}}\ , \\ W_j(d, \beta) &= \log I_{zj}^* + d(2\log\lambda_j)^* - \log^*(1 + \beta\lambda_j^{2d}) \end{aligned}$$

The elements of the Hessian matrix $H(d, \beta)$ are

$$\begin{aligned}
H_{11}(d,\beta) &= \sum_{j=1}^{m}(x_{1j}^{*})^2 - 4\beta\sum_{j=1}^{m} W_j \frac{(\log\lambda_j)^2\lambda_j^{2d}}{(1+\beta\lambda_j^{2d})^2} \\
H_{12}(d,\beta) &= \sum_{j=1}^{m} x_{1j}^{*}x_{2j}^{*} - 2\sum_{j=1}^{m} W_j \frac{(\log\lambda_j)\lambda_j^{2d}}{(1+\beta\lambda_j^{2d})^2} \\
H_{22}(d,\beta) &= \sum_{j=1}^{m}(x_{2j}^{*})^2 + \sum_{j=1}^{m} W_j \frac{\lambda_j^{4d}}{(1+\beta\lambda_j^{2d})^2}
\end{aligned}$$

Let d_0 be the true unknown memory parameter and d any admissible value and consider the same notation for the rest of parameters to estimate. Define the diagonal matrix $D_n = diag(\sqrt{m}, \lambda_m^{2d_0}\sqrt{m})$ and the matrix

$$\Omega = \begin{pmatrix} 4 & -\frac{4d_0}{(2d_0+1)^2} \\ -\frac{4d_0}{(2d_0+1)^2} & \frac{4d_0^2}{(4d_0+1)(2d_0+1)^2} \end{pmatrix}$$

Consider the following assumptions:

B.1: y_t and u_t are independent Gaussian processes.

B.2: $f_u(\lambda)$ is continuous on $[-\pi, \pi]$, bounded above and away from zero with bounded second derivative in a neighbourhood of zero.

B.3: The spectral density of y_t satisfy

$$f_y(\lambda) = C\lambda^{-2d}(1 + E\lambda^{\alpha} + o(\lambda^{\alpha}))$$

for some finite E and $\alpha \in (4d_0, 2]$.

B.4: As $n \to \infty$,

$$\frac{m^{2\alpha+1}}{n^{2\alpha}} \to K$$

for some positive constant K.

Assumption B.1 is quite severe and excludes LMSV models where u_t is not Gaussian but a log chi-square. We impose B.1 for simplicity and to directly compare our results with those in Sun and Phillips (2003). Considering recent results, Guassianity of signal and noise could be relaxed. The hypothesis of Gaussianity of y_t could be weakened as in Velasco (2000) and LMSV could also be allowed as in Deo and Hurvich (2001). However this would significantly complicate the technical details of the proofs and we prefer to keep the technical requirements to a minimum. Assumption B.2 restricts the behaviour of u_t and B.3 imposes a particular spectral behaviour of y_t around zero. As in Henry and Robinson (1996) this local specification permits to obtain the asymptotic bias of $\tilde{d}$ in terms of E. We restrict our analysis to the case $\alpha > 4d_0$ where the ALPE achieves a lower bias and higher asymptotic efficiency than the NLPE. In the standard fractional ARIMA process as considered in Sun and Phillips (2003) $\alpha = 2$. We consider also $\alpha < 2$ what may be relevant in some situations, and permit a direct extension to the seasonal or cyclical long memory case. Assumption B.4 restricts the behaviour of the bandwidth m in a similar manner as in Sun and Phillips (2003) but allowing a larger m.

Theorem 3 *Let z_t in (1) satisfy assumption B.1-B.3 and m satisfy B.4. Then as $n \to \infty$*

a) If $var(u_t) > 0$

$$D_n \begin{pmatrix} \tilde{d} - d_0 \\ \tilde{\beta} - \beta_0 \end{pmatrix} \xrightarrow{d} N\left(\Omega^{-1} b, \frac{\pi^2}{6}\Omega^{-1}\right)$$

b) If $var(u_t) = 0$

$$\sqrt{m}(\tilde{d} - d_0) \xrightarrow{d} -(\tilde{\Omega}_{11}\eta_1 + \tilde{\Omega}_{12}\eta_2)\{\tilde{\Omega}_{12}\eta_1 + \tilde{\Omega}_{22}\eta_2 \le 0\} - \Omega_{11}^{-1}\eta_1\{\tilde{\Omega}_{12}\eta_1 + \tilde{\Omega}_{22}\eta_2 > 0\}$$
$$\sqrt{m}\lambda_m^{2d_0}(\tilde{\beta} - \beta_0) \xrightarrow{d} -(\tilde{\Omega}_{12}\eta_1 + \tilde{\Omega}_{22}\eta_2)\{\tilde{\Omega}_{12}\eta_1 + \tilde{\Omega}_{22}\eta_2 \le 0\}$$

where $\tilde{\Omega} = (\tilde{\Omega}_{ij}) = \Omega^{-1}$, $\eta = (\eta_1, \eta_2)' \sim N(-b, \pi^2\Omega/6)$ and

$$b = (2\pi)^{\alpha} K 2 \begin{pmatrix} -\frac{\alpha}{(1+\alpha)^2} \\ \frac{\alpha d_0}{(2d_0+\alpha+1)(2d_0+1)(1+\alpha)} \end{pmatrix} E.$$

Sun and Phillips (2003) consider the case $y_t = (1-L)^{-d_0} w_t$ with a weak dependent w_t such that $f_z(\lambda) = (2\sin\frac{\lambda}{2})^{-2d_0}(f_w(\lambda) + (2\sin\frac{\lambda}{2})^{2d_0} f_u(\lambda))$ and then $\alpha = 2$, $C_0 = f_w(0)$, $\beta_0 = f_u(0)/f_w(0)$ and $E = \left(d_0/6 + f_w''(0)/f_w(0)\right)/2$. Whereas in Sun and Phillips (2003) the bias leading term b is different when $var(u_t) = 0$ and $var(u_t) > 0$ we do not need to discriminate both situations and in both cases the bias is of the same order.

When $var(u_t) > 0$ the asymptotic bias of $(\tilde{d}, \tilde{\beta})$ is

$$D_n^{-1}\Omega^{-1} b_n \quad = \quad D_n^{-1}\Omega^{-1}\sqrt{m}\lambda_m^{\alpha} 2 \begin{pmatrix} -\frac{\alpha}{(1+\alpha)^2} \\ \frac{\alpha d_0}{(2d_0+\alpha+1)(2d_0+1)(1+\alpha)} \end{pmatrix} E$$

$$= \frac{\lambda_m^\alpha \alpha(2d_0+1)E}{4d_0(1+\alpha)^2(2d_0+\alpha+1)} \begin{pmatrix} \alpha - 2d_0 \\ \lambda_m^{-2d_0} \frac{(2d_0+1)(4d_0+1)\alpha}{d_0} \end{pmatrix}$$

which for the processes considered in Sun and Phillips (2003) corresponds to the results in their Remark 2 but with the b_n of the $\sigma_u = 0$ case and correcting the rate of convergence in the asymptotic bias of $\tilde{\beta}$ and the $f_w(0)^2/f_u(0)^2$ term which should be $f_u(0)^2/f_w(0)^2$. In particular the asymptotic bias of $\tilde{d}$ is

$$ABias(\tilde{d}) = \left(\frac{m}{n}\right)^\alpha K_0 \quad \text{where} \quad K_0 = \frac{(2\pi)^\alpha \alpha(2d_0+1)(\alpha-2d_0)E}{4d_0(1+\alpha)^2(2d_0+\alpha+1)}$$

In contrast to the LPE and NLPE, the ALPE $\tilde{d}$ has an asymptotic positive bias which decreases with d_0. The asymptotic variance is

$$AVar(\tilde{d}) = \frac{\pi^2}{24m} C_d \quad \text{where} \quad C_d = 1 + \frac{1+4d_0}{4d_0^2}$$

and consequently the asymptotic mean squared error is

$$AMSE(\tilde{d}) = \frac{\pi^2}{24m} C_d + \left(\frac{m}{n}\right)^{2\alpha} K_0^2.$$

The optimal bandwidth, in an asymptotic MSE sense, is

$$m_{opt} = \left(\frac{\pi^2 C_d}{48\alpha K_0^2}\right)^{\frac{1}{2\alpha+1}} n^{\frac{2\alpha}{2\alpha+1}}.$$

The optimal bandwidth of the ALPE increases with n faster than the corresponding m_{opt} of the NLPE. Correspondingly the AMSE($\tilde{d}$) converges to zero at a rate $n^{-2\alpha/(2\alpha+1)}$ which is faster that the $n^{-4d_0/(4d_0+1)}$ obtained with the optimal m in the LPE and if $\alpha > 4d_0$ (as in the $\alpha = 2$ case) it is faster than the $n^{-8d_0/(8d_0+1)}$ rate achieved by the NLPE with an optimal m.

4 FINITE SAMPLE PERFORMANCE

Deo and Hurvich (2001) and Crato and Ray (2002) show that the finite sample bias of the LPE in perturbed LM series is very large, especially when the variance of the added noise is high with respect to the variance of the LM signal. A considerable bias reduction is achieved by the NLPE of Sun and Phillips (2003). We compare the finite sample performance of these two estimators with the ALPE in a LMSV

$$z_t = y_t + u_t$$

for $(1-L)^d y_t = w_t$ and $u_t = \log \varepsilon_t^2$, for ε_t and w_t independent, ε_t is standard normal and $w_t \sim N(0, \sigma_w^2)$ for $\sigma_w^2 = 0.5, 0.1$. We have chosen these low variances because they are close to the values that have been empirically found when a LMSV model is fitted to financial time series (see Breidt et al. (1998) and Pérez and Ruiz (2001) among others). These values correspond to long run noise to signal ratios $f_u(0)/f_w(0) = \pi^2$, $5\pi^2$. The first value is close to the values considered in Deo and Hurvich (2001) and Sun and Phillips (2003). The second one corresponds more closely to the values found in financial time series. We just consider a representative case $d = 0.45$. Since ε_t is standard normal, u_t is a $\log \chi_1^2$ and consequently assumption B.1 does not hold. However we consider relevant to show that the ALPE can be applied in LMSV models which are a essential tool in the modelling of financial time series, and justify in that way our conjecture of no necessity of Gaussianity of the added noise.

The Monte Carlo is carried out in SPlus 2000, generating y_t with the option arima.fracdiff.sim and for the non linear optimization we use nlminb for $0.0001 < d < 0.7$ providing the gradient and the hessian. We just consider a sample size of $n = 1024$ which is comparable with the size of many financial series and permits the exact use of the Fast Fourier Transform. The grid of bandwidths analysed is $m = 10(2)..., 500$. The number of replications is 1000.

Figure 1 shows the bias and mean square error of the LPE, NLPE and ALPE. The bias of the ALPE is significantly lower than the corresponding bias of the NLPE and the LPE and it shows a milder increase with m. The MSE of the ALPE is also more stable with m and, contrary to the LPE and NLPE, does not show a marked minimum for a small m. For a small bandwidth the MSE of the ALPE is not the lowest of the three estimation procedures but the situation adjusts as the bandwidth increases. Overall the main advantage of the ALPE is a significant reduction of the bias due to the explicit consideration of the added noise in the estimation procedure.

5 LONG MEMORY IN IBEX35 VOLATILITY

Many empirical papers have recently exposed evidence of long memory in the volatility of financial time series such as asset returns. In this section we analyze the persistence of the volatility of the Spanish stock index Ibex35 composed of the 35 more actively traded stocks.

The series covers the period 1-10-93 to 22-3-96 half-hourly. 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, $t = 1, ..., 7260$. Arteche (2003) found evidence of long memory volatility by means of the Gaussian semiparametric estimator described in equation (5) and observed that the estimates decreased with the bandwidth which could be explained by the increasing negative bias found in LMSV models. Figure 2 show the LPE, NLPE and ALPE for a grid of bandwidths $m = 6, ..., 160$. The decreasing behaviour of the LPE is similar to that of the Gaussian semiparametric estimation in Figure 3. However the NLPE and ALPE are higher and more estable with m sustaining the Monte Carlo results in the previous section.

For comparative purposes, Figure 3 shows the Gaussian semiparametric estimates (GSE) defined in (5) and the modified Gaussian semiparametric (MGSE) in equation (7). The resemblance to the LPE and ALPE is evident. The MGSE is more similar to the ALPE because contrary to the NLPE, neither of them use the linear approximation of the logarithm term.

6 CONCLUSION

The strong persistence of the volatility in many financial and economic time series and the use of LMSV models to capture such behaviour has motivated a recent interest in the estimation of the memory parameter in perturbed long memory series. The added noise gives rise to a negative bias in traditional estimators based on a local specification of the spectral density. We focus on the log periodogram regression and propose a modification that explicitly takes into account the added noise. Avoiding the log-linearization used by Sun and Phillips (2003) we attain a gain in efficiency and a significant reduction of the bias.

Although the asymptotic properties of the ALPE have been rigourously proved only under Gaussianity of signal and noise, our Monte Carlo results and other recent theoretical results in somewhat similar contexts suggest that both requirements might not be necessary, but further research following this line would be welcome.

References

[1] Andrews, D.W.K. & P. Guggenberger (2003) A bias-reduced log-periodogram regression estimator for the long-memory parameter. *Econometrica* 71, 675-712.

[2] Arteche, J. (2003) Gaussian Semiparametric Estimation in Long Memory in Stochastic Volatility and Signal Plus Noise Models. *Journal of Econometrics.* To appear.

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

[4] Arteche, J. & C. Velasco (2003) Trimming and tapering semiparametric estimates in asymmetric long memory time series. Preprint.

[5] Crato, N. & B.K. Ray (2002) Semi-parametric smoothing estimators for long-memory processes with added noise. *Journal of Statistical Planning and Inference* 105, 283-297.

[6] Breidt, F.J., N. Crato & P. de Lima (1998) The Detection and Estimation of Long Memory in Stochastic Volatility. *Journal of Econometrics* 83, 325-348.

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

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

[9] Harvey, A.C. (1998) Long memory in stochastic volatitility. In J. Knight & S. Satchell (eds.), *Forecasting Volatility in Financial Markets*, pp. 307-320. Oxford: Butterworth-Haineman.

[10] Henry, M. & P.M. Robinson (1996) Bandwidth choice in Gaussian semiparametric estimation of long range dependence. In P.M. Robinson and M. Rosenblatt (eds.), *Athens Conference on Applied Probability and Time Series, Vol. II. Lecture Notes in Statistics* 115, pp. 220-232. New York: Springer-Verlag.

[11] Hurvich, C.M. & K.I. Beltrao (1993) Asymptotics for the low-frequency ordinates of the periodogram of long-memory time series. *Journal of Time Series Analysis* 14, 455-472.

[12] Hurvich, C.M., R. Deo & J. Brodsky (1998) The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter in a long-memory time series. *Journal of Time Series Analysis* 19, 19-46.

[13] Hurvich, C.M., E. Moulines & P. Soulier (2003) Estimating long memory in volatility. Working paper SOR-2003-5, NYU Stern.

[14] Pérez, A. & E. Ruiz (2001) Finite Sample Properties of a QML Estimator of Stochastic Volatility Models with Long Memory. *Economic Letters* 70, 157-164.

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

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

[17] Robinson, P.M. & M. Henry (2003) Higher order kernel semiparametric M-estimation of long memory. *Journal of Econometrics* 114, 1-27.

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

[19] Velasco, C. (2000) Non Gaussian log-periodogram regression. *Econometric Theory* 16, 44-79.

A APPENDIX: TECHNICAL DETAILS

Proof of Theorem 1: The proof is similar to that of Theorem 1 in Hurvich and Beltrao (1993) (see also Theorem 1 in Arteche and Velasco (2003)). Write

$$L_j(d) = \int_{-n}^{n} g_n(\lambda)\mathrm{d}\lambda \tag{A.1}$$

where

$$g_n(\lambda) = K_n(\lambda_j - \lambda)\frac{f_z(\lambda)}{C\lambda_j^{-2d}}$$

and the Fejer´s kernel

$$K_n(\lambda) = \frac{1}{2\pi n} |\sum_{t=1}^{n} e^{it\lambda}|^2 = \frac{\sin^2(\frac{\lambda}{2}n)}{2\pi n \sin^2 \frac{\lambda}{2}}$$

satisfies

$$K_n(\lambda) \leq constant \times \min(n, n^{-1}\lambda^{-2}) \tag{A.2}$$

From (A.2) the integral in (A.1) over $[-\pi, -n^{-\delta}] \bigcup [n^{\delta}, \pi]$ for some $\delta \in (0, 0.5)$ is

$$O(n^{-1}|\lambda_j - n^{-\delta}|^{-2}\lambda^{2d} \int_{-\pi}^{\pi} f_z(\lambda) \mathrm{d}\lambda) = O(n^{-1}n^{2\delta}n^{-2d}) = o(n^{-2d}).$$

The integral over $(-n^{-\delta}, n^{-\delta})$ is $A_{1j} + A_{2j}$ where

$$A_{1j} = \int_{-n^{1-\delta}}^{n^{1-\delta}} \frac{\sin^2\left(\frac{2\pi j - \lambda}{2}\right)}{2\pi n^2 \sin^2\left(\frac{2\pi j - \lambda}{2n}\right)} \frac{f_y\left(\frac{\lambda}{n}\right)}{C\lambda_j^{-2d}} \mathrm{d}\lambda$$

$$A_{2j} = \int_{-n^{1-\delta}}^{n^{1-\delta}} \frac{\sin^2\left(\frac{2\pi j - \lambda}{2}\right)}{2\pi n^2 \sin^2\left(\frac{2\pi j - \lambda}{2n}\right)} \frac{f_u\left(\frac{\lambda}{n}\right)}{C\lambda_j^{-2d}} \mathrm{d}\lambda$$

and the theorem is proved letting n go to ∞. $\square$

Proof of Theorem 3: The theorem is proved as in Sun and Phillips (2003) noting that

$$x_{1j}(d, \beta) = 2\left(1 - \frac{\beta\lambda_j^{2d}}{1 + \beta\lambda_j^{2d}}\right) \log \lambda_j = 2 \log \lambda_j (1 + o(1)) \tag{A.3}$$

$$x_{2j}(d, \beta) = -\frac{\lambda_j^{2d}}{1 + \beta\lambda_j^{2d}} = -\lambda_j^{2d}(1 + o(1)) \tag{A.4}$$

which are the corresponding expressions in Sun and Phillips (2003) except a $o(1)$ term. The consistency of $\tilde{d}$ and $\tilde{\beta}$ is proved similarly noting (A.3) and (A.4). With respect to the asymptotic normality we emphasize two main differences. The first one is related with the convergence of the Hessian matrix in Lemma 5 of Sun and Phillips (2003), in particular the proof of part a),

$$\sup_{(d,\beta) \in \Theta_n} ||D_n^{-1}(H(d, \beta) - J(d, \beta))D_n^{-1}|| = o_p(1) \tag{A.5}$$

where $\Theta_n = \{(d, \beta) : |\lambda_m^{-d_0}(d - d_0)| < \varepsilon$ and $|\beta - \beta_0| < \varepsilon\}$ for $\varepsilon > 0$ arbitrary small and $J_{ab}(d, \beta) = \sum_{j=1}^{m} x_{aj}^* x_{bj}^*$, $a, b = 1, 2$. The proof that the (1,1), (1,2) and (2,1) elements of the left hand side are $o(1)$ is as in Sun and Phillips (2003) noting (A.3) and (A.4). However the (2,2) element is not zero but

$$\frac{\lambda_m^{-4d}}{m} \sum_{j=1}^{m} \frac{W_j \lambda_j^{4d}}{(1 + \beta\lambda_j^{2d})^2} = \frac{1}{m} \sum_{j=1}^{m} (a_j(d, \beta) - \bar{a}(d, \beta)) W_{1j}(d, \beta)$$

where

$$\begin{aligned}
a_j(d,\beta) &= \frac{(j/m)^{4d}}{(1+\beta\lambda_j^{2d})^2} \\
W_{1j}(d,\beta) &= V_j(d,\beta) + \epsilon_j + U_j \\
V_j(d,\beta) &= 2(d-d_0)\log\lambda_j + \log(1+\beta_0\lambda_j^{2d_0}) - \log(1+\beta\lambda_j^{2d}) \\
\epsilon_j &= \frac{\lambda_j^{\alpha}E}{1+\beta_0\lambda_j^{2d_0}} + o(\lambda_j^{\alpha}).
\end{aligned}$$

Now

$$|a_j(d,\beta)| = O\left(\left[\frac{j}{m}\right]^{4d}\right) \qquad j = 1,2,...,m,$$

and $|a_j(d,\beta) - a_{j-1}(d,\beta)|$ is bounded by

$$\begin{aligned}
&\left|\frac{(j/m)^{4d}}{(1+\beta\lambda_j^{2d})^2} - \frac{([j-1]/m)^{4d}}{(1+\beta\lambda_j^{2d})^2}\right| + \left|\frac{([j-1]/m)^{4d}}{(1+\beta\lambda_j^{2d})^2} - \frac{([j-1]/m)^{4d}}{(1+\beta\lambda_{j-1}^{2d})^2}\right| \\
= &\left|\left(\frac{j}{m}\right)^{4d}\frac{1}{(1+\beta\lambda_j^{2d})^2}\left[1-\left(\frac{j-1}{j}\right)^{4d}\right]\right| + \left|\left(\frac{j-1}{m}\right)^{4d}\frac{\beta^2(\lambda_{j-1}^{4d}-\lambda_j^{4d}) + 2\beta(\lambda_{j-1}^{2d}-\lambda_j^{2d})}{(1+\beta\lambda_j^{2d})^2(1+\beta\lambda_{j-1}^{2d})^2}\right| \\
= &\; O\left(\frac{j^{4d-1}}{m^{4d}}\right)
\end{aligned}$$

since $\lambda_{j-1}^a - \lambda_j^a = O(j^{-1}\lambda_j^a)$ for $a \neq 0$. By lemma 3 in Sun and Phillips (2003)

$$\sup_{(d,\beta)\in\Theta_n}\left|\frac{1}{m}\sum_{j=1}^{m}(a_j-\bar{a})U_j\right| = O_p\left(\frac{1}{\sqrt{m}}\right) = o_p(1)$$

Also $\sup_{(d,\beta)\in\Theta_n}\left|m^{-1}\sum_{j=1}^{m}(a_j-\bar{a})V_j(d,\beta)\right|$ is bounded by

$$\begin{aligned}
&\sup_{(d,\beta)\in\Theta_n}\left|\frac{1}{m}\sum_{j=1}^{m}(a_j-\bar{a})2(d-d_0)\log\lambda_j\right| + \sup_{(d,\beta)\in\Theta_n}\left|\frac{1}{m}\sum_{j=1}^{m}(a_j-\bar{a})\log\left(\frac{1+\beta_0\lambda_j^{2d_0}}{1+\beta\lambda_j^{2d}}\right)\right| \\
= &\; O\left(\log\lambda_m \sup_{(d,\beta)\in\Theta_n}|d-d_0|\right)) + O\left(\sup_{(d,\beta)\in\Theta_n}\lambda_m^{2d}\right) = o(1)
\end{aligned}$$

since $a_j = O(1)$, and similarly

$$\sup_{(d,\beta)\in\Theta_n}\left|\frac{1}{m}\sum_{j=1}^{m}(a_j-\bar{a})\epsilon_j\right| = O(\lambda_m^{\alpha}) = o(1)$$

and (A.5) holds. With this result the convergence of $\sup_{(d,\beta)\in\Theta_n}|D_n^{-1}H(d,\beta)D_n^{-1}|$ to Ω follows as in Sun and Phillips (2003).

The main difference with the NLPE lies on the bias. Consider

$$D_n^{-1}S(d_0,\beta_0) = \frac{1}{\sqrt{m}}\sum_{j=1}^{m} B_j(U_j+\epsilon_j)$$

where $B_j = (x_{1j}^*(d_0,\beta_0) \;\;,\;\; \lambda_m^{-2d_0}x_{2j}^*(d_0,\beta_o))'$. The bias comes from $m^{-1/2}\sum B_j\epsilon_j$ such that

$$\begin{aligned}
\frac{1}{\sqrt{m}}\sum_{j=1}^{m} x_{1j}^*(d_0,\beta_0)\epsilon_j &= \frac{2E}{\sqrt{m}}\sum_{j=1}^{m}\left(\log j - \frac{1}{m}\sum_k \log k\right)\frac{\lambda_j^\alpha}{1+\beta_0\lambda_j^{2d_0}} + o\left(\frac{m^{\alpha+\frac{1}{2}}}{n^\alpha}\right) \\
&= \frac{2E}{\sqrt{m}}\sum_{j=1}^{m}\left(\log j - \frac{1}{m}\sum_k \log k\right)\lambda_j^\alpha + o\left(\frac{m^{\alpha+\frac{1}{2}}}{n^\alpha}\right) \\
&= \frac{2E\alpha}{(1+\alpha)^2}\sqrt{m}\lambda_m^\alpha + o\left(\frac{m^{\alpha+\frac{1}{2}}}{n^\alpha}\right) \\
\frac{\lambda_m^{-2d_0}}{\sqrt{m}}\sum_{j=1}^{m} x_{2j}^*(d_0,\beta_0)\epsilon_j &= -\frac{\lambda_m^{-2d_0}E}{\sqrt{m}}\sum_{j=1}^{m}\left(\lambda_j^{2d_0} - \frac{1}{m}\sum_k \lambda_k^{2d_0}\right)\frac{\lambda_j^\alpha}{1+\beta_0\lambda_j^{2d_0}} + o\left(\frac{m^{\alpha+\frac{1}{2}}}{n^\alpha}\right) \\
&= -\frac{2d_0\alpha E}{(2d_0+\alpha+1)(2d_0+1)(1+\alpha)}\lambda_m^\alpha\sqrt{m} + o\left(\frac{m^{\alpha+\frac{1}{2}}}{n^\alpha}\right)
\end{aligned}$$

Then as $n\to\infty$

$$D_n^{-1}S(d_0,\beta_0) + b_n = \frac{1}{\sqrt{m}}\sum_{j=1}^{m} B_jU_j + o(1) \xrightarrow{d} N\left(0,\frac{\pi^2}{6}\Omega\right)$$

as in (A.34)-(A.37) in Sun and Phillips (2003). The proof when $var(u_t)=0$ follows as in their Theorem 4. $\square$

Figure 1: Bias of LPE, NLPE and ALPE, $d = 0.45$

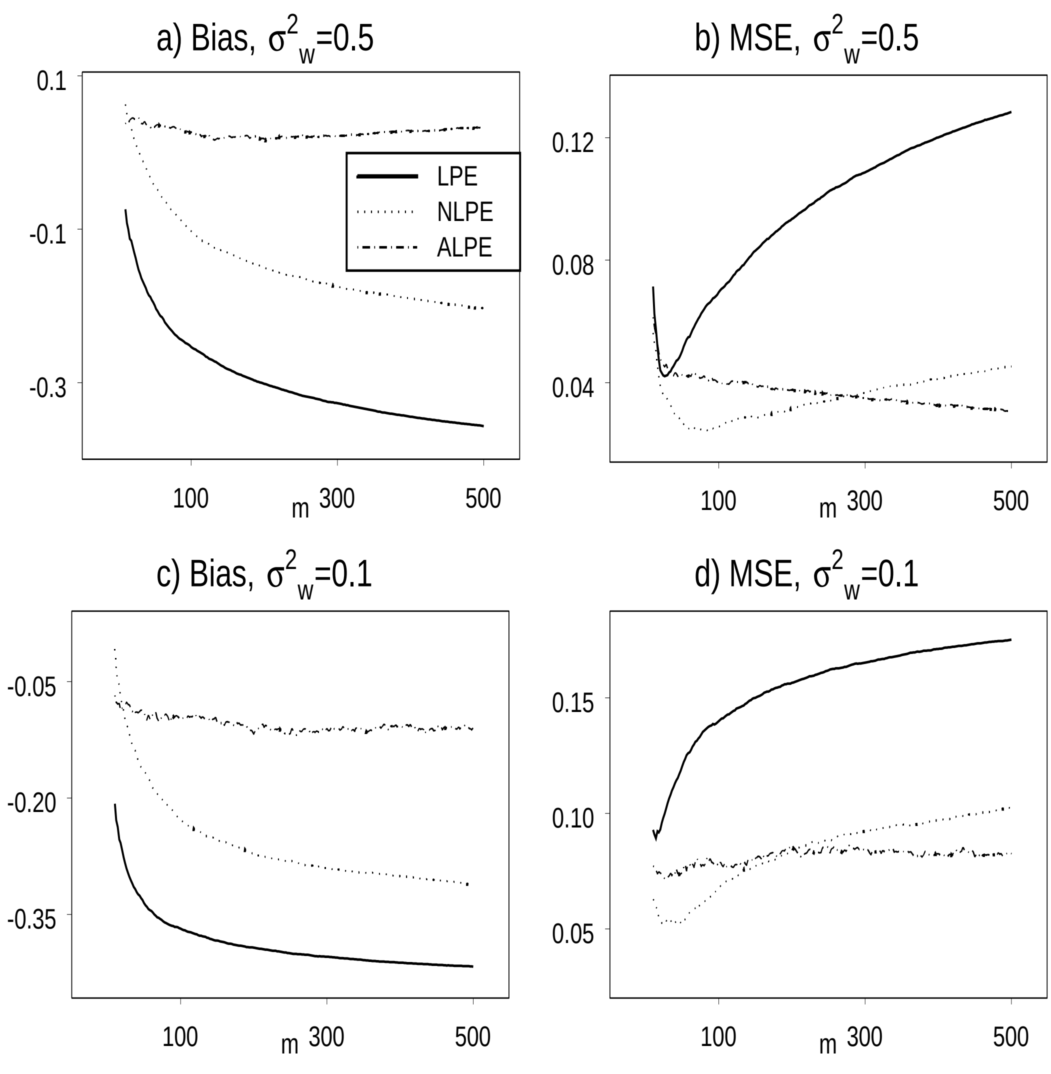

Figure 2: Log periodogram estimates (IBEX35)

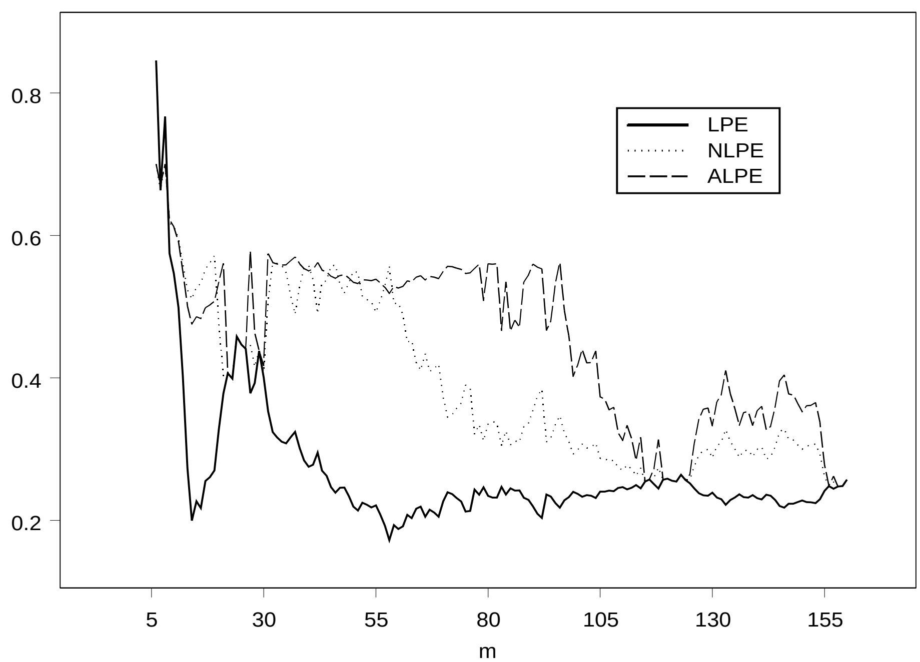

Figure 3: Standard and modified Gaussian semiparametric estimates (IBEX35)