Semiparametric estimation in perturbed long memory series

J. Arteche*†
University of the Basque Country

29th May 2005

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 in many financial and economic time series and the use of Long Memory in Stochastic Volatility (LMSV) processes to model such a behaviour. This paper discusses frequency domain semiparametric estimation of the memory parameter and proposes an extension of the log periodogram regression which explicitly accounts for the added noise, comparing it, asymptotically and in finite samples, with similar extant techniques. Contrary to the non linear log periodogram regression of Sun and Phillips (2003), we do not use a linear approximation of the logarithmic term which accounts for the added noise. A reduction of the asymptotic bias is achieved in this way and makes possible a faster convergence in long memory signal plus noise series by permitting a larger bandwidth. Monte Carlo results confirm the bias reduction but at the cost of a higher variability. An application to a series of returns of the Spanish Ibex35 stock index is finally included.

Keywords: long memory, stochastic volatility, semiparametric estimation.

*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.

† *Corresponding address*: 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: josu.arteche@ehu.es.

1 Introduction

The estimation of the memory parameter in perturbed long memory processes has recently received considerable attention motivated especially by the strong persistence found in the volatility of many financial and economic time series. Alternatively to the different extensions of ARCH and GARCH processes, the Long Memory in Stochastic Volatility (LMSV) has proved an 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 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), Sun and Phillips (2003), Arteche (2004), Hurvich et al. (2005)).

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, u_t is a weakly dependent process with a spectral density $f_u(\lambda)$ that is continuous on $[-\pi, \pi]$, bounded above and away from zero, and y_t is a long memory (LM) process characterized by a spectral density function satisfying

$$f_y(\lambda) = C\lambda^{-2d_0}(1 + O(\lambda^\alpha)) \qquad \text{as } \lambda \to 0 \tag{2}$$

for a positive finite constant C, $\alpha \in [1, 2]$ and $0 < d_0 < 0.5$. The LMSV model considers u_t a non normal white noise but in a more general signal plus noise u_t can be a serially weakly dependent process as in Sun and Phillips (2003) and Arteche (2004). The constant α is a spectral smoothness parameter which determines the adequacy of the local specification of the spectral density of y_t at frequencies around the origin. The interval $1 \leq \alpha \leq 2$ covers the most interesting situations. In parametric standard LM processes, such as the fractional ARIMA, $\alpha = 2$ and $\alpha = 1$ in the seasonal or cyclical long memory processes described in Arteche and Robinson (1999) if the long memory takes part at some frequency different from 0. The condition of positive memory $0 < d_0 < 0.5$ is usually imposed when dealing with frequency domain estimation in perturbed long memory processes and guarantees the asymptotic equivalence between spectral densities of y_t and z_t. Otherwise the memory of z_t

corresponds to that of the noise ($d_0 = 0$). For u_t uncorrelated with y_t the spectral density of z_t is

$$f_z(\lambda) = f_y(\lambda) + f_u(\lambda) = C\lambda^{-2d_0}(1 + O(\lambda^\alpha)) + f_u(\lambda) \sim C\lambda^{-2d_0}\left(1 + \frac{f_u(0)}{C}\lambda^{2d_0} + O(\lambda^\alpha)\right) \tag{3}$$

as $\lambda \to 0$ and z_t inherits the memory properties of y_t in the sense that both share the same memory parameter d_0. However the spectral smoothness parameter changes and for z_t is $\min\{2d_0, \alpha\} = 2d_0$.

The semiparametric estimators considered in this paper are based on the minimization of some function of the difference between the periodogram and the local specification of the spectral density in (3). The periodogram of z_t does not approximate accurately $C\lambda^{-2d_0}$ and this causes a bias which translates into the different estimators. This is discussed in Section 2. As a result estimation techniques have been proposed that consider explicitly the added noise in the local specification of the spectral density of z_t. They are described in Section 3. Section 4 proposes an estimator based on an extension of the log periodogram regression and establishes its asymptotic properties. Section 5 compares the "optimal" bandwidths defined as the minimizers of an approximation of the mean square error of the different semiparametric estimators considered. The performance in finite sample perturbed LM series is discussed in Section 6 by means of Monte Carlo. Section 7 shows an application to a series of returns of the Spanish Ibex35 stock index. Finally section 8 concludes. Technical details are placed in the Appendix.

2 Periodogram and local specification of the spectral density

Define

$$I_{zj} = I_z(\lambda_j) = \frac{1}{2\pi n}\left|\sum_{t=1}^{n} z_t \exp(-i\lambda_j t)\right|^2$$

the periodogram of the series z_t, $t = 1, ..., n$, at Fourier frequency $\lambda_j = 2\pi j/n$. The properties of several semiparametric estimators of d_0 depend on the adequacy of the approximation of the periodogram to the local specification of the spectral density. Hurvich and Beltrao (1993), Robinson (1995a) and Arteche and Velasco (2005) in an asymmetric long memory context, observed that the asymptotic relative bias of the periodogram produces the bias typically encountered in semiparametric estimates of the memory parameters.

Deo and Hurvich (2001), Crato and Ray (2002) and Arteche (2004) detected that the bias is quite severe in perturbed long memory series if the added noise is not explicitly considered in the estimation. It is then relevant to analyze the asymptotic bias of I_{zj} as an approximation of the local specification of the spectral density when the added noise is ignored.

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 $d_0 < 0.5$ 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*

$$L_n(j) = E\left[\frac{I_{zj}}{C\lambda_j^{-2d_0}}\right].$$

Then, considering j fixed:

$$L_n(j) = A_{1n}(j) + A_{2n}(j) + o(n^{-2d_0})$$

where

$$\lim_{n\to\infty} A_{1n}(j) = \int_{-\infty}^{\infty} \psi_j(\lambda)\left|\frac{\lambda}{2\pi j}\right|^{-2d_0} d\lambda$$

and

$$\lim_{n\to\infty} n^{2d_0} A_{2n}(j) = \int_{-\infty}^{\infty} \psi_j(\lambda)\frac{f_u(0)}{C(2\pi j)^{-2d_0}} d\lambda$$

where

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

Remark 1: The influence of the added noise turns up in $A_{2n}(j)$ and is thus asymptotically negligible if $d_0 > 0$. However for finite n $A_{2n}(j)$ can be quite large if d_0 is low and/or the long run noise to signal ratio (nsr) $f_u(0)/C$ is large. This produces the high bias of traditional semiparametric estimators which ignore the added noise in perturbed LM series and justify the modifications recently proposed and described in the next section.

Remark 2: 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.

Remark 3: When $d_0 < 0$ the bias diverges as n increases. This result was expected since the memory of z_t corresponds in this case to the memory of the noise. Then $L_n(j)$ diverges because we normalize the periodogram by a quantity that goes to zero as $n \to \infty$. As a result the estimation of a negative memory parameter of z_t is not straightforward as noted by Deo and Hurvich (2001) and Arteche (2004).

Remark 4: When $j = j(n)$ is a sequence of positive integers such that $j/n \to 0$ as $n \to \infty$, a straightforward extension of Theorem 2 in Robinson (1995a) shows that under assumptions A.1 and A.2

$$L_n(j) = 1 + O\left(\frac{\log j}{j} + \lambda_j^{\min(\alpha, 2d_0)}\right)$$

noting that

$$f_z(\lambda_j) - C\lambda_j^{-2d_0} = f_y(\lambda_j) + f_u(\lambda_j) - C\lambda_j^{-2d_0}$$

and by assumption A.1,

$$\frac{f_z(\lambda_j)}{C\lambda_j^{-2d_0}} = 1 + O\left(\lambda_j^{\min(\alpha, 2d_0)}\right).$$

3 Semiparametric estimation of the memory parameter

Let d_0 be the true unknown memory parameter and d any admissible value and consider hereafter the same notation for the rest of parameters to be estimated. The version of Robinson (1995a) of the log periodogram regression estimator (LPE), $\hat{d}_{LPE}$, is based on the least squares regression

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

where 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 (3) such that

$$\log I_{zj} = a + d_0(-2\log\lambda_j) + U_{zj} + O(\lambda_j^{2d_0}), \quad j = 1, 2, ..., m, \tag{4}$$

where $a = \log C - c$, $c = 0.577216...$ is Euler's constant and $U_{zj} = \log(I_{zj} f_z^{-1}(\lambda_j)) + c$. The bias of the least squares estimate of d_0 is dominated by the $O(\lambda_j^{2d_0})$ term which is not explicitly considered in the regression such that a negative bias of order $O(\lambda_m^{2d_0})$ arises which can be quite severe if d_0 is low. Deo and Hurvich (2001) also show that $\sqrt{m}(\hat{d}_{LPE} - d_0) \xrightarrow{d} N(0, \pi^2/24)$ as long as $m = \kappa n^{\varsigma}$ for $\varsigma < 4d_0/(4d_0 + 1)$ and κ is hereafter a generic positive constant which can be different in every case.

The main rival semiparametric estimator of the LPE is the local Whittle or Gaussian semiparametric estimator (GSE), $\hat{d}_{GSE}$, proposed by Robinson (1995b) and defined as the minimizer of

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

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). Again the bias can be approximated by a term of order $O(\lambda_m^{2d_0})$ which is caused by the added noise, and $\sqrt{m}(\hat{d}_{GSE} - d_0) \xrightarrow{d} N(0, 1/4)$ as long as $m = \kappa n^{\varsigma}$ for $\varsigma < 4d_0/(4d_0 + 1)$ (Arteche, 2004). As in the LPE, this bandwidth restriction limits quite seriously the rate of convergence of the estimators, especially if d_0 is low.

In order to reduce the bias of the GSE, Hurvich et al. (2005), noting (3), suggested to incorporate explicitly in the estimation procedure a $\beta\lambda_j^{2d}$ term which accounts for the effect of the added noise and proposed a modified Gaussian semiparametric estimator (MGSE) defined as

$$(\hat{d}_{MGSE}, \hat{\beta}_{MGSE}) = \arg\min_{\Delta\times\Theta} R(d, \beta) \tag{6}$$

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

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

When u_t is $iid(0, \sigma_u^2)$ then $f_u(\lambda) = \sigma_u^2(2\pi)^{-1}$ and $\beta_0 = \sigma_u^2(2\pi C)^{-1}$. The explicit consideration of the noise in the estimation relaxes the upper bound of the bandwidth such that $\sqrt{m}(\hat{d}_{MGSE} - d_0) \xrightarrow{d} N(0, C_{d_0}/4)$ for $C_{d_0} = 1 + (1 + 4d_0)/4d_0^2$ as long as $m = \kappa n^{\varsigma}$ for $\varsigma < 2\alpha/(2\alpha + 1)$ which permits a larger m. When $\alpha = 2$, as is typical in standard LM parametric models, $\hat{d}_{MGSE}$ achieves a rate of convergence arbitrarily close to $n^{2/5}$ which is the upper bound of the rate of convergence of $\hat{d}_{GSE}$ in the absence of additive noise. However with an additive noise the best possible rate of convergence achieved by $\hat{d}_{GSE}$ is $n^{2d_0/(4d_0+1)}$. Regarding the bias of $\hat{d}_{MGSE}$, it can be approximated by a term of order $O(\lambda_m^{\alpha})$ instead of

$O(\lambda_m^{2d_0})$ which is the order of the bias of $\hat{d}_{GSE}$ in the presence of an additive noise.

Sun and Phillips (2003) extended the log periodogram regression in a similar manner. From (3)

$$\begin{aligned}
\log I_{zj} &= \log C - c + d_0(-2\log\lambda_j) + \log\left(1 + \frac{f_u(\lambda_j)}{C}\lambda_j^{2d_0} + O(\lambda_j^{\alpha})\right) + U_{zj} \\
&= \log C - c + d_0(-2\log\lambda_j) + \log\left(1 + \frac{f_u(0)}{C}\lambda_j^{2d_0}\right) + O(\lambda_j^{\alpha}) + U_{zj} \qquad (7) \\
&= \log C - c + d_0(-2\log\lambda_j) + \frac{f_u(0)}{C}\lambda_j^{2d_0} + O(\lambda_j^{\alpha^*}) + U_{zj} \qquad (8)
\end{aligned}$$

where $\alpha^* = \min(4d_0, \alpha)$. Noting (8) Sun and Phillips (2003) proposed the following non linear regression

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

for $\beta_0 = f_u(0)/C$, such that the non linear log periodogram regression estimator (NLPE) is defined as

$$(\hat{d}_{NLPE}, \hat{\beta}_{NLPE}) = \arg\min_{\Delta\times\Theta} \sum_{j=1}^{m} (\log^* I_{zj} + d(2\log\lambda_j)^* - \beta(\lambda_j^{2d})^*)^2 \qquad (10)$$

where for a general ξ_t we use the notation $\xi_t^* = \xi_t - \bar{\xi}$ where $\bar{\xi} = \sum \xi_t/n$. The bias of $\hat{d}_{NLPE}$ is of order $O(\lambda_m^{\alpha^*})$ which is largely produced by the $O(\lambda_j^{\alpha^*})$ omitted in the regression in (9). Correspondingly $\sqrt{m}(\hat{d}_{NLPE} - d_0) \xrightarrow{d} N(0, \frac{\pi^2}{24}C_{d_0})$ as long as $m = \kappa n^{\varsigma}$ for $\varsigma < 2\alpha^*/(2\alpha^*+1)$. Sun and Phillips (2003) consider the case $\alpha = 2$ so that $\alpha^* = 4d_0$ and the behaviour of m is restricted to be $O(n^{8d_0/(8d_0+1)})$ with a bias of $\hat{d}_{NLPE}$ of order $O(\lambda_m^{4d_0})$. The upper bound of m in the NLPE is higher than in the standard LPE but lower than in the MGSE when $\alpha > 4d_0$. This is caused by the approximation of the logarithmic expression in (7). This approach has been used by Andrews and Guggenberger (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), although linear in β, is still non linear in d and the linear approximation of the logarithmic expression does not imply a significant computational advantage. Instead, noting (7) 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. We call the estimator based on a nonlinear least squares regression of (11) the augmented log periodogram regression estimator (ALPE).

4 Augmented log periodogram regression

The augmented log periodogram regression estimator (ALPE) is defined as

$$(\hat{d}_{ALPE}, \hat{\beta}_{ALPE}) = \arg\min_{\Delta\times\Theta} Q(d,\beta) \tag{12}$$

under the constraint $\beta \geq 0$, where

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

Consider the following assumptions:

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

B.2: When $var(u_t) > 0$, $f_u(\lambda)$ is continuous on $[-\pi,\pi]$, bounded above and away from zero with bounded first derivative in a neighbourhood of zero.

B.3: The spectral density of y_t satisfies

$$f_y(\lambda) = C\lambda^{-2d_0}(1 + G\lambda^{\alpha} + O(\lambda^{\alpha+\iota}))$$

for some $\iota > 0$, finite positive C, finite G, $0 < d_0 < 0.5$ and $\alpha \in (4d_0, 2] \bigcap [1,2]$.

Assumption B.1 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). Assumption B.2 restricts the behaviour of u_t as in Assumption 1 in Sun and Phillips (2003). Assumption B.3 imposes a particular spectral behaviour of y_t around zero relaxing Assumption 2 in Sun and Phillips (2003). As in Henry and Robinson (1996) this local specification permits to obtain the leading part of the asymptotic bias of $\hat{d}_{ALPE}$ in terms of G. We restrict our analysis to the case $\alpha > 4d_0$ where the ALPE achieves a lower bias and higher asymptotic efficiency than the NLPE by permitting a larger m. When $\alpha \leq 4d_0$ the ALPE and the NLPE share the same asymptotic distribution with the same upper bound of m. In the standard fractional ARIMA process considered in Sun and Phillips (2003) $\alpha = 2 > 4d_0$.

Theorem 2 *Under assumptions B.1-B.3, as* $n \to \infty$ $\hat{d}_{ALPE} - d_0 = o_p(1)$ *if* $1/m + m/n \to 0$, *and* $\hat{d}_{ALPE} - d_0 = O_p((m/n)^{2d_0})$, $\hat{\beta}_{ALPE} - \beta_0 = o_p(1)$ *if* $m/n + n^{4d_0(1+\delta)}/m^{4d_0(1+\delta)+1} \to 0$ *for some arbitrary small* $\delta > 0$.

This is the same result as the consistency of the NLPE in Theorem 2 in Sun and Phillips (2003) and can be proved similarly noting that

$$\frac{1}{m}Q(d,\beta)=\frac{1}{m}\sum_{j=1}^{m}\left\{U_{zj}^{*}+V_{j}^{*}+O(\lambda_j^{\alpha})\right\}^2$$

for $V_j^*=V_j^*(d,\beta)=V_j(d,\beta)-\bar{V}(d,\beta)$, $V_j(d,\beta)=2(d-d_0)\log\lambda_j+\log(1+\beta_0\lambda_j^{2d_0})-\log(1+\beta\lambda_j^{2d})$ and that $\log(1+\beta\lambda_j^{2d})=\beta\lambda_j^{2d}+O(\lambda_j^{4d})$ for $(d,\beta)\in\Delta\times\Theta$.

The main difference of the ALPE with respect to the NLPE lies in the asymptotic distribution, particularly in the term responsible of the asymptotic bias. The first order conditions of the minimization problem are

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

where Λ is the Lagrange multiplier pertaining to the constraints $\beta\geq 0$ and

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

with

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

The Hessian matrix $H(d,\beta)$ has elements

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

Define $D_n=diag(\sqrt{m},\lambda_m^{2d_0}\sqrt{m})$ and the matrix

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

and consider the following assumptions

B.4: d_0 is an interior point of Δ and $0 \leq \beta_0 < \Theta_1$.

B.5: As $n \to \infty$,

$$\frac{m^{\alpha+0.5}}{n^{\alpha}} \to K$$

for some positive constant K.

The structure of the series, if perturbed or not, is not known beforehand. It is then interesting to consider not only the case $var(u_t) > 0$ but also the no added noise case, $var(u_t) = 0$, and analyze the behaviour of the ALPE in both situations.

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

a) If $var(u_t) > 0$

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

b) If $var(u_t) = 0$

$$\sqrt{m}(\hat{d}_{ALPE} - d_0) \xrightarrow{d} -(\tilde{\Omega}_{11}\eta_1 + \tilde{\Omega}_{12}\eta_2)\{\tilde{\Omega}_{12}\eta_1 + \tilde{\Omega}_{22}\eta_2 \leq 0\} - \Omega_{11}^{-1}\eta_1\{\tilde{\Omega}_{12}\eta_1 + \tilde{\Omega}_{22}\eta_2 > 0\}$$

$$\sqrt{m}\lambda_m^{2d_0}(\hat{\beta}_{ALPE} - \beta_0) \xrightarrow{d} -(\tilde{\Omega}_{12}\eta_1 + \tilde{\Omega}_{22}\eta_2)\{\tilde{\Omega}_{12}\eta_1 + \tilde{\Omega}_{22}\eta_2 \leq 0\}$$

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

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

Sun and Phillips (2003) consider $y_t = (1-L)^{-d_0} w_t$ 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 $\alpha = 2$, $C = f_w(0)$, $\beta_0 = f_u(0)/f_w(0)$ and $G = \left(d_0/6 + f_w''(0)/f_w(0)\right)/2$. Whereas in Sun and Phillips (2003) the term leading the asymptotic bias, 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 asymptotic bias is of the same order. To eliminate this bias we have to choose a bandwidth of order $o(n^{\alpha/(\alpha+0.5)})$ instead of that in assumption B.5.

When $var(u_t) > 0$ the asymptotic bias of $(\hat{d}_{ALPE}, \hat{\beta}_{ALPE})$ can be approximated by

$$\begin{aligned} D_n^{-1}\Omega^{-1}b_n &= D_n^{-1}\Omega^{-1}\sqrt{m}\lambda_m^{\alpha} 2 \begin{pmatrix} -\frac{\alpha}{(1+\alpha)^2} \\ \frac{\alpha d_0}{(2d_0+\alpha+1)(2d_0+1)(1+\alpha)} \end{pmatrix} G \\ &= \frac{\lambda_m^{\alpha}\alpha(2d_0+1)G}{4d_0(1+\alpha)^2(2d_0+\alpha+1)} \begin{pmatrix} \alpha - 2d_0 \\ \lambda_m^{-2d_0}\frac{(2d_0+1)(4d_0+1)\alpha}{d_0} \end{pmatrix} \end{aligned}$$

which for the processes considered in Sun and Phillips (2003) corresponds to the result in their Remark 2 but with the b_n of their $\sigma_u = 0$ case and correcting the rate of convergence in the asymptotic bias of $\hat{\beta}_{NLPE}$ and the $f_w(0)^2/f_u(0)^2$ term which should be $f_u(0)^2/f_w(0)^2$ in their formula (48). The asymptotic bias of $\hat{d}_{ALPE}$ can then be approximated by

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

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

$$AVar(\hat{d}_{ALPE}) = \frac{\pi^2}{24m} C_{d_0}$$

and consequently the asymptotic mean squared error can be approximated by

$$AMSE(\hat{d}_{ALPE}) = \frac{\pi^2}{24m} C_{d_0} + \left(\frac{m}{n}\right)^{2\alpha} K_0^2.$$

5 Comparing "optimal" bandwidths

The role of the bandwidth on semiparametric memory parameter estimates is crucial to get reliable estimates. A too large choice of m can induce a large bias whereas a too small m generates a high variability of the estimates. An optimal choice of m is usually obtained minimizing an approximate form of the mean square error (MSE). In this section we compare the optimal bandwidths obtained in this way for the estimators considered above in the long memory signal plus noise process characterized by assumptions B.1-B.3 with $\sigma_u^2 > 0$.

By Sun and Phillips (2003, Theorem 1), the asymptotic bias of $\hat{d}_{LPE}$ can be approximated by

$$ABias(\hat{d}_{LPE}) = -\beta_0 \frac{d_0}{(2d_0+1)^2} \lambda_m^{2d_0} \tag{13}$$

and considering the asymptotic variance $\pi^2/(24m)$ the bandwidth that minimizes the approximate MSE is

$$m_{LPE}^{opt} = \left[\frac{\pi^2}{24} \frac{(2d_0+1)^4}{(2\pi)^{4d_0}\beta_0^2 4d_0^3}\right]^{\frac{1}{4d_0+1}} n^{\frac{4d_0}{4d_0+1}} \tag{14}$$

Using similar arguments to those employed by Henry and Robinson (1996) it is easy to show that the asymptotic bias of $\hat{d}_{GSE}$ can also be approximated by (13). In consequence the optimal bandwidth is (Arteche, 2004)

$$m_{GSE}^{opt} = \left[\frac{1}{4} \frac{(2d_0+1)^4}{(2\pi)^{4d_0}\beta_0^2 4d_0^3}\right]^{\frac{1}{4d_0+1}} n^{\frac{4d_0}{4d_0+1}} = \left(\frac{AVar(\hat{d}_{GSE})}{AVar(\hat{d}_{LPE})}\right)^{\frac{1}{4d_0+1}} m_{LPE}^{opt} \tag{15}$$

and since $AVar(\hat{d}_{GSE}) < AVar(\hat{d}_{LPE})$ then $m^{opt}_{GSE} < m^{opt}_{LPE}$.

Similarly the optimal bandwidth of the NLPE is given by Sun and Phillips (2003)

$$m^{opt}_{NLPE} = \left[\frac{\pi^2 C_{d_0}(4d_0+1)^4(6d_0+1)^2}{192 d_0^3(2d_0+1)^2(2\pi)^{8d_0}\beta_0^4}\right]^{\frac{1}{8d_0+1}} n^{\frac{8d_0}{8d_0+1}} \tag{16}$$

The ALPE share the same asymptotic variance as the NLPE but the lower order bias produces a higher optimal bandwidth. Minimizing $AMSE(\hat{d}_{ALPE})$ the optimal bandwidth is

$$m^{opt}_{ALPE} = \left(\frac{\pi^2 C_{d_0}}{48\alpha K_0^2}\right)^{\frac{1}{2\alpha+1}} n^{\frac{2\alpha}{2\alpha+1}}.$$

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

The ALPE is comparable in terms of optimal bandwidth and bias with the MGSE. In fact, using similar similar arguments to those suggested by Henry and Robinson (1996) it is straightforward to show that the bias of $\hat{d}_{MGSE}$ can be approximated by that of $\hat{d}_{ALPE}$

$$ABias(\hat{d}_{MGSE}) = ABias(\hat{d}_{ALPE}) = \left(\frac{m}{n}\right)^{\alpha} K_0$$

and then

$$m^{opt}_{MGSE} = \left(\frac{AVar(\hat{d}_{MGSE})}{AVar(\hat{d}_{ALPE})}\right)^{\frac{1}{4d_0+1}} m^{opt}_{ALPE} = \left(\frac{C_{d_0}}{8\alpha K_0^2}\right)^{\frac{1}{2\alpha+1}} n^{\frac{2\alpha}{2\alpha+1}}.$$

Contrary to $\hat{d}_{LPE}$, $\hat{d}_{GSE}$ and $\hat{d}_{NLPE}$, the asymptotic bias of $\hat{d}_{ALPE}$ and $\hat{d}_{MGSE}$ do not depend on β_0 and consequently m^{opt}_{ALPE} and m^{opt}_{MGSE} are invariant to different values of nsr $f_u(0)/C$.

6 Finite sample performance

Deo and Hurvich (2001), Crato and Ray (2002) and Arteche (2004) have shown that the bias in perturbed LM series of $\hat{d}_{LPE}$ and $\hat{d}_{GSE}$ is very high and increases considerably with m, especially when the nsr is large. Consequently a very low bandwidth should be used to get reliable estimates, at least in terms of bias. A substantial bias reduction is achieved by including the added noise explicitly in the estimation procedure as in $\hat{d}_{NLPE}$, $\hat{d}_{ALPE}$ and $\hat{d}_{MGSE}$. We compare the finite sample performance of these estimators in a LMSV

$$z_t = y_t + u_t$$

for $(1-L)^{d_0} y_t = w_t$ 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 (e.g. Breidt et al. (1998), Pérez and Ruiz (2001)). These values correspond to long run nsr $f_u(0)/f_w(0) = \pi^2$, $5\pi^2$. The first one is close to the ratios considered in Deo and Hurvich (2001), Sun and Phillips (2003) and Hurvich and Ray (2003). The second corresponds more closely to the values found in financial time series. We consider $d_0 = 0.2$, 0.45 and 0.8. For $d_0 = 0.8$ the process is not stationary and is even larger than 0.75 so that the proof of the asymptotic normality of $\hat{d}_{MGSE}$ in Hurvich et al. (2005) does not apply. However the estimators are expected to perform well as long as $d_0 < 1$ (Sun and Phillips, 2003). Also, since ε_t is standard normal, u_t is a $\log \chi_1^2$ and assumption B.1 does not hold. However we consider relevant to show that these estimators can be applied in LMSV models which are an 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 over 1000 replications in SPlus 2000, generating y_t with the option arima.fracdiff.sim and for the different non linear optimizations we use nlminb for $0.01 < d < 1$ and $\exp(-20) < \beta < \exp(8)$ providing the gradient and the hessian. We consider sample sizes $n = 1024$, 4096 and 8192 which are comparable with the size of many financial series and permits the exact use of the Fast Fourier Transform. For each sample size we take four different bandwidths $m = [n^{0.4}]$, $[n^{0.6}]$, $[n^{0.8}]$ and m_{est}^{opt} for $est = LPE$, $NLPE$, $ALPE$, GSE and $MGSE$ with the constraint $5 \leq m_{est}^{opt} \leq [n/2-1]$. Table 1 displays m_{est}^{opt} for the different values of d_0, n and σ_w^2. The lower constraint applies for the LPE and GSE for low d_0 and/or σ_w^2 and also for the NLPE for $d_0 = 0.2$ and $\sigma_w^2 = 0.1$. The upper limit is applicable for the ALPE and MGSE with the lower sample size. Note that m_{ALPE}^{opt} and m_{MGSE}^{opt} do not depend on the nsr.

TABLES 1 AND 2 ABOUT HERE

Table 2 shows the bias and MSE of the estimators across the models considered. The following conclusions can be deduced:

- The bias of the LPE and GSE is very high, especially for a large bandwidth and nsr. The bias clearly reduces with the estimation techniques which account for the added noise.

- In terms of bias the NLPE tends to be overcome by the ALPE and MGSE especially for the high nsr case. The bias of the ALPE and MGSE is more invariant to different values of the nsr and more stable with the bandwidth while a large choice of m produces an extremely high bias of the NLPE. The NLPE tends to beat both ALPE and MGSE in terms of MSE for an appropriate choice of m and low values of d_0. In any other case $\hat{d}_{ALPE}$ and $\hat{d}_{MGSE}$ are better choices.

- Regarding the behaviour of the different estimators using the "optimal" bandwidth, the best performance in terms of MSE corresponds to the MGSE which has the lowest MSE in 16 out of 18 cases, followed by the ALPE which has lower MSE than the NLPE, GSE and LPE in 13 out of 18 cases. Only for $d_0 = 0.2$ and $d_0 = 0.45$ with $n = 1024$ the ALPE is overwhelmed by the LPE, GSE or NLPE. It deserves special mention the situation for $d_0 = 0.2$ and $n = 1024$ since here the LPE and GSE are the best choices. This was somehow expected because for such a low value of d there is not much scope for bias and also the estimates are constrained to be larger than 0.01 limiting the size of the bias. For $d_0 = 0.45, 0.85$ the MGSE and the ALPE have a lower MSE than the LPE, GSE and NLPE (only for $d_0 = 0.45$, $n = 1024$ and $\sigma_w^2 = 0.5$ the NLPE has a lower MSE than the ALPE).

- The "optimal" bandwidth performs better than the other three bandwidths for the ALPE and MGSE suggesting that a large m should be chosen. However the NLPE tends to have lower MSE with $m = n^{0.6}$ in those cases where $n^{0.6}$ is larger than m_{NLPE}^{opt} which occurs in every case when $n = 1024$, and for $n = 4096$ and $n = 8192$ except when $d_0 = 0.8$ and $\sigma_w^2 = 0.5$, suggesting that m_{NLPE}^{opt} tends to be undervalued.

We also compute the coverage probabilities of the nominal 90% confidence intervals obtained with the five estimators using the asymptotic normality of all of them (although this is not true for $d_0 = 0.8$ we keep the normality assumption for comparative purposes). For each we use two different standard errors. First we use the variance in the asymptotic distributions. For $\hat{d}_{LPE}$ and $\hat{d}_{GSE}$ these are $\pi^2/(24m)$ and $1/(4m)$. The rest of estimators have asymptotic variances which depend on the unknown memory parameter d_0, $(1+2d_0)^2/(16d_0^2m)$ for $\hat{d}_{MGSE}$ and $\pi^2(1+2d_0)^2/(96d_0^2m)$ for $\hat{d}_{NLPE}$ and $\hat{d}_{ALPE}$. To get feasible expressions we substitute the unknown d_0 with the corresponding estimates. We also use the finite sample hessian based approximations for the standard errors suggested by Deo and Hurvich (2001), Hurvich and Ray (2003) and Sun and Phillips (2003). For $\hat{d}_{LPE}$, $\hat{d}_{GSE}$ and $\hat{d}_{ALPE}$ these are

$$\begin{aligned}
\widehat{var}(\hat{d}_{LPE}) &= \frac{\pi^2}{24}\left(\sum_{j=1}^{m}\left(\log\lambda_j - \frac{1}{m}\sum_{k=1}^{m}\log\lambda_k\right)^2\right)^{-1} \\
\widehat{var}(\hat{d}_{GSE}) &= \left(4\sum_{j=1}^{m}\left(\log\lambda_j - \frac{1}{m}\sum_{k=1}^{m}\log\lambda_k\right)^2\right)^{-1} \\
\widehat{var}(\hat{d}_{ALPE}) &= SE_J + (SE_H - SE_J)I(H(\hat{d}_{ALPE},\hat{\beta}_{ALPE}) > 0) \\
SE_H &= \frac{\pi^2}{6}\frac{H_{22}(\hat{d}_{ALPE},\hat{\beta}_{ALPE})}{H_{11}(\hat{d}_{ALPE},\hat{\beta}_{ALPE})H_{22}(\hat{d}_{ALPE},\hat{\beta}_{ALPE}) - H_{12}(\hat{d}_{ALPE},\hat{\beta}_{ALPE})^2} \\
SE_J &= \frac{\pi^2}{6}\frac{J_{n,22}(\hat{d}_{ALPE},\hat{\beta}_{ALPE})}{J_{n,11}(\hat{d}_{ALPE},\hat{\beta}_{ALPE})J_{n,22}(\hat{d}_{ALPE},\hat{\beta}_{ALPE}) - J_{n,12}(\hat{d}_{ALPE},\hat{\beta}_{ALPE})^2}
\end{aligned}$$

where $I(H(\hat{d}_{ALPE},\hat{\beta}_{ALPE}) > 0) = 1$ if $H(\hat{d}_{ALPE},\hat{\beta}_{ALPE})$ is positive definite and 0 otherwise and $J_n(d,\beta)$ is defined in the proof of Theorem 3. $\widehat{var}(\hat{d}_{NLPE})$ is similarly obtained as defined in formulae (60) and (61) in Sun and Phillips (2003). We have also tried only SE_J and while this approach performs significantly worse in the NLPE it renders slightly worse ALPE confidence intervals for low m and n and similar for large values of the bandwidth and sample size. $\widehat{var}(\hat{d}_{MGSE})$ is defined in formula (16) in Hurvich and Ray (2003)[1] with the unknowns substituted with the corresponding estimates.

TABLES 3, 4 AND 5 ABOUT HERE

Tables 3, 4 and 5 display the coverage frequencies, mean and median lengths of the 90% Gaussian based confidence intervals on $d_0 = 0.2$, 0.45 and 0.8 respectively, constructed using the asymptotic variances with estimated d_0 (Prob.A, Mean.A and Med.A) and the finite sample hessian approximation (Prob.H, Mean.H and Med.H). The following comment deserve particular attention:

- The coverage frequencies of the LPE and GSE are satisfactory only for a low bandwidth but as m increases they go rapidly towards zero. Here mean and median lengths are

[1]Note that $b_{1,0}^{-1}$ in formula (16) of Hurvich and Ray (2003) corresponds to β_0 in our notation.

equal because the approximations used for the standard errors do not depend on estimates and do not vary across simulations. The finite sample approximation of the standard error tends to give wider intervals and better (closer to the nominal 90%) coverage frequencies.

- The NLPE has close to nominal coverage frequencies for $d_0 = 0.2$ but as d_0, n and m increase the frequencies go down, being close to zero in several situations ($d_0 = 0.45$, $m = n^{0.8}$, $n = 4096$, 8192, and $d_0 = 0.8$, $m = n^{0.8}$ for all n) . For $d_0 = 0.2$ the finite sample approximation of the standard error tends to give narrower intervals and better coverage than the feasible asymptotic expression. However as d_0 increases the situation changes and for $d_0 = 0.8$ the asymptotic expression gives in many cases better coverage even with narrower intervals.

- For $d_0 = 0.2$ the performance of the confidence intervals based on ALPE and MGSE is quite poor with very wide intervals and with mean lengths much higher than the median, especially for low m and n. This fact was also noted by Hurvich and Ray (2003) and explained by the existence of outlying estimates of d_0. The intervals based on the finite sample approximation of the standard errors can be extremely wide, especially with a large nsr, due to large variations in the estimated nsr that require larger sample sizes and bandwidths to be accurately estimated. For higher values of d_0 and large n the ALPE and MGSE confidence intervals behave significantly better when the finite sample approximation of the standard error is used. Overall the MGSE confidence intervals tend to perform better than the intervals based on ALPE.

- Comparing the different estimators there is not one that outperforms the others in every situation and the best choice depends on n, m, d_0 and the nsr. Overall the NLPE seems a good choice for low d_0 and n but for values of d_0 close to the stationary limit or higher and a large sample size the MGSE (and the ALPE) with the finite sample approximated standard error is a wiser choice.

7 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 a series of returns 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, $t = 1, ..., 7260$. We use as the proxy of the volatility the series $y_t = \log(x_t - \bar{x})^2$ which corresponds to the volatility component in a LMSV model apart from an added noise. Arteche (2004) found evidence of long memory in y_t by means of the GSE and observed that the estimates decreased rapidly with the bandwidth which could be explained by the increasing negative bias of the GSE found in LMSV models.

Figure 1 shows the LPE, GSE, NLPE, MGSE and ALPE for a grid of bandwidths $m = 25, ..., 300$ together with the 95% confidence intervals obtained using both the feasible asymptotic expression and the finite sample approximations of the standard errors described in Section 6. We do not consider higher values of m to avoid distorting influence of seasonality. To elude the phenomenon encountered in the Monte Carlo of excessively wide intervals we restrict the values of the standard errors to be lower than an arbitrary value of 0.6 such that if it exceeds that value we take the standard error calculated with a bandwidth increased by one. This situation only occurs with $\hat{d}_{NLPE}$ for $m = 29$ when the approximated standard error is 3.03. Both approximations of the standard errors provide similar intervals for the LPE and GSE and for most of the bandwidths also for the NLPE. Only very low values of m lead to significant different intervals. The situation is different for the MGSE and ALPE where the finite sample approximations always give wider intervals, especially for low values of m.

It is also observable that the LPE and GSE decrease with m faster than the other estimates. This situation is more clearly displayed in Figure 2 which shows the five estimates for a grid of bandwidth $m = 25, ..., 200$. The LPE and GSE behave similarly with a rapid decrease with m. This can be due to a large negative bias caused by some unaccounted for added noise. In this situations a sensible strategy is to estimate d by techniques that account for the added noise such as the NLPE, ALPE or MGSE because the large bias of the LPE and GSE can render these estimates meaningless. The NLPE remains high for a wider range of values of m but finally decreases for lower values of m than the MGSE and ALPE which behave quite similarly. This is consistent with the asymptotic and finite sample results described in the previous sections.

Finally Figure 3 shows estimates and confidence intervals for $m = 150, ..., 300$. The GSE and LPE give strong support in favour of the stationarity of the volatility. However the NLPE, ALPE and MGSE cast some doubt about it, at least with a 95% confidence. Taking into account the results described in the previous sections, we should be cautious in concluding in favour of the stationarity of the volatility of this series of Ibex35 returns.

FIGURES 1, 2 AND 3 ABOUT HERE

8 CONCLUSION

The strong persistence of the volatility in many financial and economic time series and the use of LMSV models to capture such a 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 which can be reduced by including explicitly the added noise in the estimation procedure as the NLPE and MGSE. We have proposed an additional log periodogram regression based estimator, the ALPE, whose properties are close to those of the MGSE, which seems the better option in a wide range of possibilities. In particular both show a significant improvement in terms of bias but at the cost of a larger finite sample variance than the NLPE for low values of d, bandwidth and sample size. However, for large sample sizes and high values of d the ALPE and MGSE perform significantly better than the NLPE, especially if the nsr is large as is often the case in financial time series.

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 (2005)). Write

$$L_n(j) = \int_{-n}^{n} g_{nj}(\lambda) \mathrm{d}\lambda \tag{A.1}$$

where

$$g_{nj}(\lambda) = K_n(\lambda_j - \lambda)\frac{f_z(\lambda)}{C\lambda_j^{-2d_0}} , \quad K_n(\lambda) = \frac{1}{2\pi n}|\sum_{t=1}^{n} e^{it\lambda}|^2 = \frac{\sin^2(\frac{\lambda}{2}n)}{2\pi n \sin^2 \frac{\lambda}{2}}$$

and the Fejer´s kernel 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

$$O(n^{-1}|\lambda_j - n^{-\delta}|^{-2}\lambda^{2d_0} \int_{-\pi}^{\pi} f_z(\lambda)\mathrm{d}\lambda) = O(n^{-1}n^{2\delta}n^{-2d_0}) = o(n^{-2d_0}).$$

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

$$A_{1n}(j) = \int_{-n^{1-\delta}}^{n^{1-\delta}} \frac{\sin^2\left(\frac{2\pi j - \lambda}{2}\right)}{2\pi n^2 \sin^2\left(\frac{2\pi j-\lambda}{2n}\right)} \frac{f_y\left(\frac{\lambda}{n}\right)}{C\lambda_j^{-2d_0}}\mathrm{d}\lambda$$

$$A_{2n}(j) = \int_{-n^{1-\delta}}^{n^{1-\delta}} \frac{\sin^2\left(\frac{2\pi j - \lambda}{2}\right)}{2\pi n^2 \sin^2\left(\frac{2\pi j-\lambda}{2n}\right)} \frac{f_u\left(\frac{\lambda}{n}\right)}{C\lambda_j^{-2d_0}}\mathrm{d}\lambda$$

and the theorem is proved letting n go to ∞. □

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

$$x_{1j}(d,\beta) = 2\left(1 - \frac{\beta\lambda_j^{2d}}{1+\beta\lambda_j^{2d}}\right)\log\lambda_j = 2\log\lambda_j(1-\beta\lambda_j^{2d}) + O(\lambda_j^{4d}\log\lambda_j) \tag{A.3}$$

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

for $(d, \beta) \in \Delta \times \Theta$. This approximation leads to two main differences in the proof of the asymptotic normality. Noting the consistency of $\hat{d}_{ALPE}$ the first one is related to the convergence of the Hessian matrix in Lemma 5 of Sun and Phillips (2003), in particular the proof of part a),

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

where $\Theta_n = \{(d,\beta) : |\lambda_m^{-d_0}(d - d_0)| < \varepsilon \text{ and } |\beta - \beta_0| < \varepsilon\}$ for $\varepsilon > 0$ arbitrary small and $J_{n,ab}(d,\beta) = \sum_{j=1}^m x_{aj}^* x_{bj}^*$, $a, b = 1, 2$. 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

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

where

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

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)

$$\sup_{(d,\beta)\in\Theta_n}\left|\frac{1}{m}\sum_{j=1}^m a_j^*(d,\beta)U_{zj}\right| = O_p\left(\frac{1}{\sqrt{m}}\right) = o_p(1)$$

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

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

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

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

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) noting (A.3) and (A.4).

The second difference with the NLPE lies on the bias term. Consider

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

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

$$\begin{aligned}
\frac{1}{\sqrt{m}} \sum_{j=1}^{m} x^*_{1j}(d_0, \beta_0)\epsilon_j &= \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \left[(1-\beta_0 \lambda_j^{2d_0}) \log \lambda_j\right]^* \left(\frac{G\lambda_j^{\alpha}}{1+\beta_0 \lambda_j^{2d_0}} + O\left(\lambda_j^{\alpha+\iota}\right)\right) \\
&+ O(\sqrt{m}\lambda_m^{4d_0+\alpha} \log \lambda_m) \\
&= \frac{2}{\sqrt{m}} \sum_{j=1}^{m} \log^* \lambda_j \left(G\lambda_j^{\alpha} + O(\lambda_j^{2d_0+\alpha})\right) + O(\sqrt{m}\lambda_m^{\alpha+\iota} \log \lambda_m) \\
&+ O(\sqrt{m}\lambda_m^{2d_0+\alpha} \log \lambda_m) + O(\sqrt{m}\lambda_m^{4d_0+\alpha} \log \lambda_m) \\
&= \frac{2G}{\sqrt{m}} \sum_{j=1}^{m} \left(\log j - \frac{1}{m} \sum_k \log k \right) \lambda_j^{\alpha} + o\left(\sqrt{m}\lambda_m^{\alpha}\right) \\
&= \frac{2G\alpha}{(1+\alpha)^2} \sqrt{m}\lambda_m^{\alpha}(1+o(1)) \\
\frac{\lambda_m^{-2d_0}}{\sqrt{m}} \sum_{j=1}^{m} x^*_{2j}(d_0, \beta_0)\epsilon_j &= -\frac{\lambda_m^{-2d_0} G}{\sqrt{m}} \sum_{j=1}^{m} \left(\lambda_j^{2d_0} - \frac{1}{m} \sum_k \lambda_k^{2d_0} \right) \frac{\lambda_j^{\alpha}}{1+\beta_0 \lambda_j^{2d_0}} + O\left(\sqrt{m}\lambda_m^{\alpha+\iota}\right) \\
&= -\frac{2d_0 \alpha G}{(2d_0+\alpha+1)(2d_0+1)(1+\alpha)} \lambda_m^{\alpha} \sqrt{m}(1+o(1))
\end{aligned}$$

Then as $n \to \infty$

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

as in (A.34)-(A.37) in Sun and Phillips (2003) with minor modifications to adapt their proofs to our assumption B.1-B.3. Since the rest of the proof relies heavily on Sun and Phillips (2003) and Robinson (1995a) we omit the details. The proof when $var(u_t) = 0$ follows as in Theorem 4 in Sun and Phillips (2003). □

References

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

[2] Arteche, J., 2004, Gaussian Semiparametric Estimation in Long Memory in Stochastic Volatility and Signal Plus Noise Models. *J. Econometrics* 119, 131-154.

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

[4] Arteche, J., Velasco, C., 2005. Trimming and tapering semiparametric estimates in asymmetric long memory time series. *Journal of Time Series Analysis*, forthcoming.

[5] Crato, N., Ray, B.K., 2002. Semi-parametric smoothing estimators for long-memory processes with added noise. *J. Statist. Plann. Inference* 105, 283-297.

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

[7] Deo, R.S., Hurvich, C.M., 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., Porter-Hudak, S., 1983. The estimation and application of long-memory time series models. *J. Time Ser. Anal.* 4, 221-238.

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

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

[11] Hurvich, C.M., Beltrao, K.I., 1993. Asymptotics for the low-frequency ordinates of the periodogram of long-memory time series. *J. Time Ser. Anal.* 14, 455-472.

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

[13] Hurvich, C.M., Moulines, E., Soulier, P., 2005. Estimating long memory in volatility. *Econometrica*, forthcoming.

[14] Hurvich, C.M., Ray, B.K., 2003, The Local Whittle Estimator of Long-Memory Stochastic Volatility. *Journal of Financial Econometrics* 1, 445-470.

[15] Pérez, A., Ruiz, E., 2001. Finite Sample Properties of a QML Estimator of Stochastic Volatility Models with Long Memory. *Econ. Lett.* 70, 157-164.

[16] Robinson, P.M., 1995a. Log-periodogram regression of time series with long-range dependence. *Ann. Statist.* 23, 1048-1072.

[17] Robinson, P.M., 1995b. Gaussian semiparametric estimation of long-range dependence. *Ann. Statist.* 23, 1630-1661.

[18] Robinson, P.M., Henry, M., 2003. Higher order kernel semiparametric M-estimation of long memory. *J. Econometrics* 114, 1-27.

[19] Sun, Y., Phillips, P.C.B., 2003. Nonlinear log-periodogram regression for perturbed fractional processes. *J. Econometrics* 115, 355-389.

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

Table 1: "Optimal" bandwidths

		$\sigma_w^2 = 0.5$					$\sigma_w^2 = 0.1$				
n	d_0	LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE
1024	0.2	6	5	12	511	511	5	5	5	511	511
	0.45	13	11	29	511	502	5	5	7	511	502
	0.8	27	24	53	511	511	12	11	22	511	511
4096	0.2	12	9	29	1895	1715	5	5	5	1895	1715
	0.45	32	27	87	1681	1522	10	8	21	1681	1522
	0.8	79	70	177	2047	1936	36	32	74	2047	1936
8192	0.2	16	12	45	3299	2987	5	5	5	3299	2987
	0.45	51	42	149	2927	2650	16	13	36	2927	2650
	0.8	134	119	323	3723	3370	62	55	135	3723	3370

Table 2: Bias and MSE

			$m=n^{0.4}$					$m=n^{0.6}$					$m=n^{0.8}$					$m=m_{est}^{opt}$				
n	σ_w^2		LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE
			$d_0=0.2$																			
1024	0.5	Bias	-0.075	-0.101	0.110	0.111	0.042	-0.132	-0.144	-0.013	0.078	0.037	-0.163	-0.167	-0.080	0.049	0.028	-0.005	-0.015	0.149	0.043	0.022
		MSE	0.025	0.023	0.073	0.128	0.097	0.022	0.023	0.019	0.109	0.092	0.027	0.028	0.014	0.091	0.081	0.057	0.058	0.102	0.079	0.067
	0.1	Bias	-0.103	-0.127	0.090	0.054	-0.001	-0.152	-0.164	-0.041	-0.004	-0.037	-0.175	-0.180	-0.102	-0.032	-0.058	0.018	-0.029	0.288	-0.032	-0.061
		MSE	0.025	0.025	0.071	0.110	0.083	0.026	0.028	0.020	0.082	0.072	0.031	0.033	0.017	0.070	0.064	0.075	0.054	0.228	0.070	0.063
4096	0.5	Bias	-0.093	-0.106	0.076	0.105	0.053	-0.137	-0.143	-0.047	0.069	0.033	-0.162	-0.164	-0.093	0.055	0.033	-0.046	-0.062	0.062	0.031	0.016
		MSE	0.019	0.019	0.050	0.093	0.069	0.021	0.022	0.013	0.072	0.058	0.027	0.027	0.013	0.057	0.044	0.031	0.032	0.043	0.041	0.030
	0.1	Bias	-0.125	-0.142	0.039	0.016	-0.027	-0.165	-0.170	-0.076	-0.017	-0.041	-0.182	-0.184	-0.125	-0.023	-0.036	0.015	-0.026	0.199	-0.054	-0.069
		MSE	0.023	0.025	0.042	0.069	0.057	0.029	0.030	0.016	0.062	0.052	0.033	0.034	0.019	0.055	0.051	0.076	0.059	0.150	0.047	0.043
8192	0.5	Bias	-0.091	-0.105	0.060	0.098	0.056	-0.136	-0.139	-0.055	0.066	0.037	-0.162	-0.162	-0.092	0.037	0.006	-0.057	-0.064	0.033	0.022	-0.001
		MSE	0.017	0.017	0.037	0.075	0.059	0.020	0.020	0.012	0.060	0.045	0.026	0.026	0.012	0.034	0.022	0.024	0.026	0.027	0.025	0.015
	0.1	Bias	-0.132	-0.146	0.009	0.003	-0.034	-0.169	-0.173	-0.090	-0.012	-0.045	-0.184	-0.185	-0.138	-0.045	-0.050	0.026	-0.014	0.183	-0.043	-0.041
		MSE	0.023	0.025	0.029	0.057	0.047	0.029	0.031	0.016	0.055	0.046	0.034	0.034	0.022	0.045	0.042	0.079	0.063	0.131	0.041	0.037
			$d_0=0.45$																			
1024	0.5	Bias	-0.107	-0.132	0.081	0.162	0.091	-0.222	-0.225	-0.065	0.111	0.042	-0.321	-0.315	-0.173	0.062	0.012	-0.105	-0.123	0.000	0.049	0.019
		MSE	0.049	0.047	0.075	0.127	0.103	0.057	0.056	0.024	0.095	0.072	0.105	0.100	0.039	0.062	0.042	0.056	0.059	0.041	0.048	0.030
	0.1	Bias	-0.247	-0.276	-0.031	0.033	-0.045	-0.344	-0.354	-0.192	-0.007	-0.045	-0.400	-0.401	-0.286	-0.039	-0.061	-0.096	-0.141	0.122	-0.028	-0.032
		MSE	0.088	0.096	0.071	0.131	0.118	0.124	0.129	0.060	0.129	0.119	0.161	0.162	0.090	0.117	0.107	0.121	0.114	0.140	0.116	0.103
4096	0.5	Bias	-0.054	-0.073	0.081	0.130	0.071	-0.165	-0.166	-0.042	0.045	0.008	-0.294	-0.281	-0.160	0.016	0.005	-0.069	-0.076	-0.011	0.017	0.009
		MSE	0.025	0.021	0.049	0.069	0.051	0.031	0.030	0.012	0.031	0.021	0.087	0.079	0.034	0.012	0.007	0.022	0.020	0.014	0.008	0.005
	0.1	Bias	-0.181	-0.198	0.002	0.070	0.017	-0.310	-0.308	-0.157	0.051	0.009	-0.392	-0.385	-0.258	0.033	0.007	-0.094	-0.122	0.023	0.028	0.007
		MSE	0.051	0.053	0.045	0.072	0.063	0.100	0.097	0.037	0.063	0.051	0.154	0.148	0.070	0.048	0.033	0.074	0.079	0.059	0.040	0.027
8192	0.5	Bias	-0.039	-0.050	0.087	0.127	0.078	-0.141	-0.140	-0.033	0.023	0.003	-0.282	-0.267	-0.127	0.012	0.004	-0.053	-0.055	-0.012	0.010	0.005
		MSE	0.018	0.013	0.038	0.053	0.037	0.022	0.021	0.008	0.017	0.011	0.080	0.072	0.018	0.006	0.003	0.012	0.011	0.008	0.004	0.002
	0.1	Bias	-0.154	-0.161	-0.001	0.058	0.019	-0.284	-0.280	-0.133	0.028	0.002	-0.383	-0.372	-0.240	0.032	0.010	-0.085	-0.097	0.008	0.027	0.010
		MSE	0.039	0.037	0.035	0.051	0.043	0.083	0.080	0.026	0.036	0.028	0.147	0.139	0.059	0.021	0.013	0.047	0.048	0.033	0.018	0.011
			$d_0=0.8$																			
1024	0.5	Bias	-0.019	-0.031	0.057	0.070	0.043	-0.141	-0.148	-0.007	0.032	0.007	-0.410	-0.372	-0.174	0.041	0.024	-0.037	-0.039	0.010	0.039	0.026
		MSE	0.035	0.028	0.037	0.038	0.033	0.030	0.029	0.017	0.023	0.019	0.171	0.141	0.035	0.016	0.011	0.024	0.018	0.018	0.014	0.010
	0.1	Bias	-0.110	-0.138	0.027	0.050	0.001	-0.344	-0.339	-0.124	0.021	-0.009	-0.589	-0.541	-0.356	0.026	0.009	-0.082	-0.095	0.001	0.032	0.014
		MSE	0.055	0.053	0.042	0.045	0.043	0.129	0.123	0.035	0.035	0.033	0.349	0.296	0.137	0.026	0.020	0.057	0.053	0.035	0.025	0.020
4096	0.5	Bias	0.016	0.007	0.074	0.083	0.062	-0.060	-0.065	0.023	0.041	0.021	-0.352	-0.307	-0.132	0.030	0.021	-0.010	-0.012	0.011	0.028	0.021
		MSE	0.021	0.015	0.026	0.027	0.021	0.008	0.007	0.008	0.011	0.007	0.125	0.096	0.019	0.005	0.003	0.008	0.006	0.007	0.004	0.003
	0.1	Bias	-0.015	-0.032	0.055	0.063	0.036	-0.216	-0.210	-0.044	0.035	0.021	-0.538	-0.464	-0.294	0.032	0.019	-0.029	-0.035	0.007	0.031	0.024
		MSE	0.022	0.016	0.026	0.028	0.021	0.052	0.048	0.010	0.015	0.011	0.290	0.217	0.088	0.010	0.006	0.018	0.014	0.014	0.008	0.006
8192	0.5	Bias	0.025	0.014	0.078	0.083	0.062	-0.036	-0.039	0.025	0.035	0.022	-0.321	-0.279	-0.112	0.025	0.020	-0.006	-0.008	0.012	0.025	0.019
		MSE	0.014	0.009	0.020	0.021	0.015	0.004	0.003	0.006	0.007	0.005	0.104	0.079	0.014	0.003	0.002	0.004	0.003	0.004	0.003	0.002
	0.1	Bias	0.012	-0.003	0.071	0.075	0.051	-0.164	-0.159	-0.020	0.031	0.021	-0.514	-0.433	-0.271	0.027	0.019	-0.012	-0.015	0.016	0.029	0.022
		MSE	0.015	0.011	0.021	0.021	0.016	0.030	0.028	0.006	0.010	0.006	0.265	0.188	0.075	0.006	0.003	0.010	0.008	0.009	0.005	0.003

Table 3: 90% Confidence Intervals ($d_0 = 0.2$)

		$m = n^{0.4}$					$m = n^{0.6}$					$m = n^{0.8}$					$m = m_{est}^{opt}$				
σ_w^2		LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE
												$n = 1024$									
0.5	Prob.A	0.977	0.982	0.906	0.764	0.798	0.396	0.208	0.983	0.739	0.768	0.002	0.000	0.998	0.709	0.728	0.930	0.906	0.882	0.673	0.693
	Mean.A	0.527	0.411	6.508	11.13	9.769	0.264	0.206	3.225	5.047	4.048	0.132	0.103	1.713	2.412	1.885	0.861	0.736	7.334	1.631	1.241
	Med.A	0.527	0.411	1.451	2.168	3.050	0.264	0.206	1.013	1.283	1.463	0.132	0.103	0.748	0.898	0.888	0.861	0.736	1.592	0.513	0.470
	Prob.H	0.994	0.993	0.924	0.985	1.000	0.476	0.254	0.947	0.964	1.000	0.003	0.000	0.886	0.966	0.998	0.986	0.980	0.904	0.974	0.999
	Mean.H	0.690	0.538	1.878	82783.5	63447.0	0.294	0.229	0.817	19971.3	14448.7	0.137	0.107	0.477	5791.7	3820.4	1.424	1.294	2.087	3093.3	1988.4
	Med.H	0.690	0.538	1.319	6.082	11.56	0.294	0.229	0.645	3.199	4.433	0.137	0.107	0.379	2.126	2.360	1.424	1.294	1.576	1.582	1.737
0.1	Prob.A	0.988	0.993	0.902	0.800	0.831	0.287	0.084	0.988	0.824	0.829	0.002	0.000	1.000	0.801	0.832	0.915	0.909	0.772	0.785	0.817
	Mean.A	0.527	0.411	7.322	13.498	11.53	0.264	0.206	4.144	6.946	5.840	0.132	0.103	2.455	3.723	3.183	0.944	0.736	9.905	2.724	2.294
	Med.A	0.527	0.411	1.581	5.116	16.83	0.264	0.206	1.241	4.345	5.406	0.132	0.103	0.972	3.264	5.243	0.944	0.736	1.937	3.372	3.505
	Prob.H	0.998	0.995	0.899	0.984	1.000	0.340	0.119	0.962	0.980	1.000	0.002	0.000	0.843	0.985	1.000	1.000	0.985	0.907	0.984	1.000
	Mean.H	0.690	0.538	1.726	101888.4	77328.3	0.294	0.229	0.822	28820.7	23462.4	0.137	0.107	0.476	10282.9	9016.2	1.660	1.294	5.916	6794.8	5265.9
	Med.H	0.690	0.538	1.306	14.901	51.87	0.294	0.229	0.675	8.338	13.805	0.137	0.107	0.383	4.899	10.456	1.660	1.294	3.841	4.488	5.050
												$n = 4096$									
0.5	Prob.A	0.991	0.601	0.900	0.696	0.742	0.157	0.034	0.990	0.686	0.718	0.000	0.000	1.000	0.629	0.642	0.951	0.950	0.930	0.641	0.643
	Mean.A	0.406	0.317	4.554	7.160	5.550	0.174	0.136	1.954	2.350	1.805	0.076	0.059	0.803	0.815	0.541	0.609	0.548	4.380	0.498	0.359
	Med.A	0.406	0.317	1.231	1.399	1.412	0.174	0.136	0.786	0.746	0.748	0.076	0.059	0.432	0.259	0.222	0.609	0.548	1.175	0.171	0.142
	Prob.H	0.995	0.997	0.889	0.956	1.000	0.185	0.044	0.935	0.955	1.000	0.000	0.000	0.806	0.940	0.990	0.986	0.985	0.914	0.962	0.984
	Mean.H	0.492	0.383	1.392	38375.2	25984.9	0.185	0.144	0.630	5768.3	3845.2	0.077	0.060	0.296	776.0	276.6	0.842	0.809	1.258	304.06	94.355
	Med.H	0.492	0.383	1.008	3.396	4.959	0.185	0.144	0.474	1.751	1.970	0.077	0.060	0.247	1.037	0.917	0.842	0.809	0.982	0.763	0.657
0.1	Prob.A	0.996	0.405	0.933	0.802	0.818	0.037	0.007	0.998	0.778	0.801	0.000	0.000	1.000	0.733	0.753	0.908	0.903	0.969	0.770	0.787
	Mean.A	0.406	0.317	5.878	10.11	8.640	0.174	0.136	2.847	4.403	3.515	0.076	0.059	1.432	1.966	1.562	0.944	0.736	10.31	1.389	1.179
	Med.A	0.406	0.317	1.389	3.597	6.523	0.174	0.136	0.988	2.345	2.199	0.076	0.059	0.674	1.533	1.403	0.944	0.736	2.137	1.519	1.483
	Prob.H	0.999	1.000	0.911	0.974	1.000	0.047	0.007	0.927	0.983	1.000	0.000	0.000	0.744	0.978	1.000	1.000	0.974	0.896	0.983	1.000
	Mean.H	0.492	0.383	1.295	57661.7	45030.0	0.185	0.144	0.591	13555.5	11057.0	0.077	0.060	0.332	3942.6	2938.7	1.660	1.294	8.354	2416.8	1967.8
	Med.H	0.492	0.383	0.993	8.628	19.74	0.185	0.144	0.466	3.977	4.913	0.077	0.060	0.267	2.360	2.731	1.660	1.294	4.052	2.211	2.557
												$n = 8192$									
0.5	Prob.A	0.694	0.560	0.939	0.705	0.725	0.062	0.005	0.995	0.666	0.698	0.000	0.000	1.000	0.616	0.676	0.967	0.953	0.949	0.614	0.508
	Mean.A	0.352	0.274	3.721	5.425	4.226	0.142	0.110	1.425	1.530	1.023	0.057	0.045	0.495	0.426	0.306	0.527	0.475	2.987	0.271	0.198
	Med.A	0.352	0.274	1.071	1.110	1.068	0.142	0.110	0.658	0.571	0.484	0.057	0.045	0.310	0.190	0.160	0.527	0.475	1.020	0.124	0.106
	Prob.H	0.998	0.651	0.908	0.950	1.000	0.071	0.008	0.919	0.941	0.996	0.000	0.000	0.823	0.952	0.980	0.985	0.980	0.934	0.955	0.969
	Mean.H	0.412	0.321	1.157	24013.3	16070.3	0.148	0.116	0.579	2629.3	1019.3	0.058	0.045	0.236	79.33	15.132	0.690	0.656	1.073	30.482	10.043
	Med.H	0.412	0.321	0.887	2.474	3.687	0.148	0.116	0.406	1.271	1.262	0.058	0.045	0.194	0.716	0.575	0.690	0.656	0.834	0.524	0.436
0.1	Prob.A	0.525	0.331	0.970	0.804	0.825	0.007	0.000	0.999	0.740	0.786	0.000	0.000	1.000	0.748	0.750	0.899	0.894	0.965	0.717	0.652
	Mean.A	0.352	0.274	5.083	8.360	6.906	0.142	0.110	2.332	3.328	2.739	0.057	0.045	1.267	1.534	1.177	0.944	0.736	9.559	0.970	0.736
	Med.A	0.352	0.274	1.300	2.773	3.086	0.142	0.110	0.891	1.654	1.811	0.057	0.045	0.662	1.319	1.081	0.944	0.736	2.077	0.933	0.369
	Prob.H	1.000	0.437	0.946	0.985	1.000	0.008	0.000	0.901	0.980	1.000	0.000	0.000	0.581	0.979	0.999	1.000	0.970	0.929	0.973	0.998
	Mean.H	0.412	0.321	1.234	41115.5	31034.2	0.148	0.116	0.540	8644.6	6411.0	0.058	0.045	0.295	2580.1	1860.5	1.660	1.294	10.06	1362.8	925.67
	Med.H	0.412	0.321	0.875	5.970	8.663	0.148	0.116	0.415	2.843	3.701	0.058	0.045	0.233	2.124	2.007	1.660	1.294	4.424	1.586	1.536

Prob.A, Mean.A, Med.A denote coverage frequencies, mean lengths and median lengths of the nominal 90% confidence intervals with the asymptotic expression for standard errors with estimated parameters. Prob.H, Mean.H, Med.H denote coverage frequencies, mean lengths and median lengths of the nominal 90% confidence intervals with the finite sample Hessian based approximation of the standard errors with estimated parameters.

Table 4: 90% Confidence Intervals ($d_0 = 0.45$)

		$m = n^{0.4}$					$m = n^{0.6}$					$m = n^{0.8}$					$m = m_{est}^{opt}$				
σ_w^2		LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE
											$n = 1024$										
0.5	Prob.A	0.747	0.625	0.881	0.677	0.710	0.153	0.037	0.988	0.632	0.717	0.000	0.000	0.724	0.515	0.443	0.745	0.653	0.944	0.393	0.382
	Mean.A	0.527	0.411	2.236	2.658	2.002	0.264	0.206	0.721	0.685	0.540	0.132	0.103	0.550	0.324	0.257	0.585	0.496	1.324	0.225	0.171
	Med.A	0.527	0.411	1.022	0.940	0.808	0.264	0.206	0.598	0.510	0.442	0.132	0.103	0.358	0.266	0.217	0.585	0.496	0.828	0.190	0.153
	Prob.H	0.864	0.776	0.922	0.995	1.000	0.192	0.055	0.943	0.984	1.000	0.000	0.000	0.415	0.977	0.939	0.882	0.807	0.923	0.956	0.934
	Mean.H	0.690	0.538	1.824	12658.2	6947.8	0.294	0.229	0.637	342.48	73.813	0.137	0.107	0.397	26.769	21.098	0.796	0.698	1.139	17.577	0.563
	Med.H	0.690	0.538	1.399	2.304	2.540	0.294	0.229	0.572	1.285	1.046	0.137	0.107	0.278	0.776	0.616	0.796	0.698	0.963	0.630	0.507
0.1	Prob.A	0.491	0.306	0.948	0.767	0.765	0.005	0.000	0.997	0.658	0.678	0.000	0.000	0.513	0.612	0.481	0.906	0.550	1.000	0.462	0.384
	Mean.A	0.527	0.411	3.821	6.019	5.017	0.264	0.206	1.933	2.826	1.965	0.132	0.103	1.086	1.486	1.021	0.944	0.736	5.956	1.066	0.721
	Med.A	0.527	0.411	1.174	1.125	1.080	0.264	0.206	0.780	0.638	0.549	0.132	0.103	0.511	0.309	0.254	0.944	0.736	1.458	0.217	0.165
	Prob.H	0.647	0.440	0.956	0.989	1.000	0.010	0.000	0.874	0.990	1.000	0.000	0.000	0.090	0.985	0.999	1.000	1.000	0.999	0.986	0.985
	Mean.H	0.690	0.538	1.835	39960.7	27831.6	0.294	0.229	0.735	9319.8	5207.6	0.137	0.107	0.424	3333.6	1786.2	1.660	1.294	3.457	2201.4	1267.4
	Med.H	0.690	0.538	1.363	3.093	4.597	0.294	0.229	0.613	1.960	2.468	0.137	0.107	0.338	1.582	1.618	1.660	1.294	2.593	1.408	1.492
											$n = 4096$										
0.5	Prob.A	0.805	0.750	0.884	0.745	0.769	0.085	0.015	0.985	0.742	0.748	0.000	0.000	0.219	0.559	0.534	0.781	0.724	0.966	0.486	0.494
	Mean.A	0.406	0.317	1.037	1.106	0.756	0.174	0.136	0.404	0.376	0.301	0.076	0.059	0.482	0.162	0.127	0.373	0.317	0.524	0.109	0.089
	Med.A	0.406	0.317	0.793	0.755	0.634	0.174	0.136	0.384	0.351	0.287	0.076	0.059	0.197	0.157	0.124	0.373	0.317	0.487	0.107	0.088
	Prob.H	0.886	0.822	0.899	0.927	0.997	0.110	0.021	0.956	0.960	0.959	0.000	0.000	0.110	0.905	0.896	0.864	0.813	0.944	0.903	0.893
	Mean.H	0.492	0.383	1.317	1415.0	263.90	0.185	0.144	0.395	0.683	0.499	0.077	0.060	0.297	0.353	0.272	0.443	0.383	0.589	0.291	0.230
	Med.H	0.492	0.383	1.072	1.415	1.290	0.185	0.144	0.364	0.620	0.484	0.077	0.060	0.148	0.345	0.269	0.443	0.383	0.509	0.283	0.227
0.1	Prob.A	0.564	0.373	0.945	0.811	0.756	0.000	0.000	1.000	0.606	0.558	0.000	0.000	0.028	0.324	0.285	0.718	0.631	0.918	0.252	0.234
	Mean.A	0.406	0.317	1.528	1.890	1.309	0.174	0.136	0.670	0.589	0.399	0.076	0.059	0.316	0.234	0.149	0.667	0.582	1.848	0.145	0.102
	Med.A	0.406	0.317	0.851	0.773	0.661	0.174	0.136	0.457	0.344	0.284	0.076	0.059	0.266	0.154	0.125	0.667	0.582	0.948	0.107	0.089
	Prob.H	0.688	0.483	0.928	0.958	1.000	0.000	0.000	0.850	0.940	0.997	0.000	0.000	0.006	0.931	0.944	0.984	0.976	0.914	0.906	0.917
	Mean.H	0.492	0.383	1.321	6247.8	2912.0	0.185	0.144	0.464	499.19	146.53	0.077	0.060	0.205	79.495	0.619	0.959	0.884	1.508	13.249	0.535
	Med.H	0.492	0.383	0.990	1.637	1.870	0.185	0.144	0.376	0.937	0.874	0.077	0.060	0.177	0.678	0.557	0.959	0.884	1.143	0.588	0.483
											$n = 8192$										
0.5	Prob.A	0.815	0.773	0.887	0.749	0.767	0.051	0.010	0.976	0.793	0.759	0.000	0.000	0.037	0.578	0.578	0.812	0.774	0.983	0.492	0.520
	Mean.A	0.352	0.274	0.790	0.769	0.574	0.142	0.110	0.319	0.303	0.239	0.057	0.045	0.166	0.121	0.095	0.295	0.254	0.380	0.082	0.067
	Med.A	0.352	0.274	0.674	0.650	0.546	0.142	0.110	0.311	0.294	0.233	0.057	0.045	0.145	0.119	0.094	0.295	0.254	0.368	0.081	0.067
	Prob.H	0.883	0.832	0.893	0.918	0.991	0.064	0.011	0.958	0.950	0.939	0.000	0.000	0.011	0.903	0.894	0.872	0.843	0.968	0.910	0.906
	Mean.H	0.412	0.321	1.032	105.44	1.065	0.148	0.116	0.310	0.478	0.353	0.058	0.045	0.119	0.241	0.187	0.335	0.293	0.416	0.199	0.159
	Med.H	0.412	0.321	0.855	1.111	0.989	0.148	0.116	0.290	0.449	0.349	0.058	0.045	0.108	0.238	0.186	0.335	0.293	0.363	0.197	0.158
0.1	Prob.A	0.587	0.417	0.947	0.857	0.781	0.000	0.000	0.931	0.653	0.523	0.000	0.000	0.001	0.375	0.333	0.762	0.683	0.940	0.271	0.265
	Mean.A	0.352	0.274	1.160	1.191	0.815	0.142	0.110	0.405	0.331	0.260	0.057	0.045	0.202	0.125	0.097	0.527	0.456	0.984	0.086	0.069
	Med.A	0.352	0.274	0.738	0.680	0.580	0.142	0.110	0.361	0.290	0.231	0.057	0.045	0.192	0.117	0.094	0.527	0.456	0.734	0.080	0.067
	Prob.H	0.668	0.506	0.933	0.956	1.000	0.000	0.000	0.710	0.943	0.945	0.000	0.000	0.000	0.932	0.920	0.866	0.822	0.932	0.919	0.899
	Mean.H	0.412	0.321	1.119	1865.9	654.26	0.148	0.116	0.336	0.743	0.589	0.058	0.045	0.133	0.487	0.367	0.690	0.620	1.131	6.468	0.330
	Med.H	0.412	0.321	0.856	1.304	1.276	0.148	0.116	0.292	0.682	0.559	0.058	0.045	0.126	0.443	0.351	0.690	0.620	0.834	0.396	0.317

Prob.A, Mean.A, Med.A denote coverage frequencies, mean lengths and median lengths of the nominal 90% confidence intervals with the asymptotic expression for standard errors with estimated parameters. Prob.H, Mean.H, Med.H denote coverage frequencies, mean lengths and median lengths of the nominal 90% confidence intervals with the finite sample Hessian based approximation of the standard errors with estimated parameters.

Table 5: 90% Confidence Intervals ($d_0 = 0.8$)

		$m = n^{0.4}$					$m = n^{0.6}$					$m = n^{0.8}$					$m = m_{est}^{opt}$				
σ_w^2		LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE	LPE	GSE	NLPE	ALPE	MGSE
												$n = 1024$									
0.5	Prob.A	0.891	0.859	1.000	1.000	1.000	0.448	0.294	0.900	0.759	0.731	0.000	0.000	0.182	0.547	0.543	0.863	0.768	0.979	0.436	0.444
	Mean.A	0.527	0.411	0.895	0.890	0.676	0.264	0.206	0.435	0.428	0.337	0.132	0.103	0.239	0.212	0.166	0.406	0.336	0.475	0.150	0.117
	Med.A	0.527	0.411	0.809	0.796	0.640	0.264	0.206	0.426	0.418	0.331	0.132	0.103	0.237	0.210	0.165	0.406	0.336	0.467	0.148	0.117
	Prob.H	0.940	0.911	0.996	1.000	1.000	0.511	0.343	0.975	0.998	0.965	0.000	0.000	0.125	0.945	0.884	0.921	0.897	0.994	0.900	0.862
	Mean.H	0.690	0.538	1.522	231.83	1.270	0.294	0.229	0.496	0.670	0.504	0.137	0.107	0.200	0.426	0.330	0.492	0.413	0.579	0.380	0.296
	Med.H	0.690	0.538	1.289	1.503	1.184	0.294	0.229	0.464	0.639	0.499	0.137	0.107	0.198	0.415	0.328	0.492	0.413	0.532	0.370	0.292
0.1	Prob.A	0.787	0.657	1.000	1.000	1.000	0.023	0.006	0.829	0.593	0.569	0.000	0.000	0.024	0.414	0.393	0.841	0.783	1.000	0.276	0.289
	Mean.A	0.527	0.411	0.971	1.012	0.727	0.264	0.206	0.470	0.438	0.345	0.132	0.103	0.430	0.216	0.169	0.609	0.496	0.776	0.152	0.119
	Med.A	0.527	0.411	0.822	0.801	0.652	0.264	0.206	0.456	0.415	0.331	0.132	0.103	0.277	0.210	0.166	0.609	0.496	0.720	0.148	0.117
	Prob.H	0.867	0.773	0.992	1.000	1.000	0.027	0.007	0.851	0.986	0.930	0.000	0.000	0.023	0.949	0.948	0.920	0.877	0.994	0.952	0.929
	Mean.H	0.690	0.538	1.503	1152.43	160.502	0.294	0.229	0.473	0.923	0.701	0.137	0.107	0.302	0.697	0.535	0.842	0.698	1.133	0.644	0.499
	Med.H	0.690	0.538	1.278	1.649	1.354	0.294	0.229	0.446	0.845	0.678	0.137	0.107	0.210	0.648	0.517	0.842	0.698	0.974	0.605	0.484
												$n = 4096$									
0.5	Prob.A	0.923	0.778	1.000	1.000	0.991	0.636	0.499	0.865	0.763	0.809	0.000	0.000	0.062	0.592	0.594	0.823	0.824	0.862	0.436	0.440
	Mean.A	0.406	0.317	0.647	0.644	0.505	0.174	0.136	0.281	0.279	0.219	0.076	0.059	0.133	0.122	0.095	0.237	0.197	0.257	0.075	0.060
	Med.A	0.406	0.317	0.627	0.623	0.496	0.174	0.136	0.279	0.277	0.218	0.076	0.059	0.133	0.121	0.095	0.237	0.197	0.256	0.075	0.060
	Prob.H	0.946	0.867	1.000	1.000	1.000	0.674	0.532	0.893	0.896	0.895	0.000	0.000	0.040	0.841	0.831	0.867	0.858	0.879	0.809	0.807
	Mean.H	0.492	0.383	1.004	1.048	0.778	0.185	0.144	0.305	0.368	0.272	0.077	0.060	0.110	0.204	0.158	0.261	0.218	0.269	0.175	0.137
	Med.H	0.492	0.383	0.874	0.957	0.757	0.185	0.144	0.292	0.351	0.271	0.077	0.060	0.109	0.203	0.158	0.261	0.218	0.263	0.173	0.137
0.1	Prob.A	0.893	0.800	1.000	1.000	0.985	0.043	0.012	0.865	0.695	0.690	0.000	0.000	0.000	0.438	0.446	0.813	0.788	0.895	0.322	0.319
	Mean.A	0.406	0.317	0.655	0.653	0.513	0.174	0.136	0.291	0.281	0.220	0.076	0.059	0.151	0.122	0.095	0.352	0.291	0.401	0.075	0.060
	Med.A	0.406	0.317	0.634	0.630	0.502	0.174	0.136	0.289	0.277	0.218	0.076	0.059	0.151	0.121	0.095	0.352	0.291	0.396	0.075	0.060
	Prob.H	0.922	0.872	0.999	1.000	1.000	0.051	0.015	0.861	0.947	0.884	0.000	0.000	0.000	0.880	0.874	0.909	0.853	0.944	0.865	0.851
	Mean.H	0.492	0.383	1.008	1.103	0.804	0.185	0.144	0.282	0.429	0.332	0.077	0.060	0.112	0.303	0.234	0.412	0.345	0.489	0.278	0.217
	Med.H	0.492	0.383	0.883	0.984	0.784	0.185	0.144	0.276	0.419	0.330	0.077	0.060	0.112	0.299	0.233	0.412	0.345	0.432	0.274	0.215
												$n = 8192$									
0.5	Prob.A	0.835	0.827	0.999	0.999	0.997	0.718	0.591	0.859	0.799	0.813	0.000	0.000	0.042	0.593	0.575	0.846	0.840	0.879	0.448	0.447
	Mean.A	0.352	0.274	0.556	0.555	0.436	0.142	0.110	0.228	0.227	0.178	0.057	0.045	0.099	0.092	0.072	0.182	0.151	0.190	0.056	0.046
	Med.A	0.352	0.274	0.547	0.545	0.432	0.142	0.110	0.227	0.226	0.178	0.057	0.045	0.099	0.092	0.072	0.182	0.151	0.190	0.056	0.046
	Prob.H	0.973	0.890	0.999	0.997	1.000	0.740	0.622	0.876	0.882	0.881	0.000	0.000	0.033	0.809	0.776	0.867	0.862	0.883	0.782	0.792
	Mean.H	0.412	0.321	0.888	0.831	0.622	0.148	0.116	0.246	0.286	0.208	0.058	0.045	0.082	0.144	0.112	0.195	0.162	0.191	0.123	0.097
	Med.H	0.412	0.321	0.722	0.769	0.615	0.148	0.116	0.236	0.270	0.207	0.058	0.045	0.082	0.144	0.112	0.195	0.162	0.189	0.122	0.097
0.1	Prob.A	0.825	0.796	0.998	0.998	0.994	0.067	0.023	0.872	0.725	0.721	0.000	0.000	0.000	0.440	0.452	0.807	0.782	0.877	0.306	0.341
	Mean.A	0.352	0.274	0.559	0.558	0.438	0.142	0.110	0.233	0.228	0.178	0.057	0.045	0.112	0.092	0.072	0.268	0.222	0.294	0.056	0.046
	Med.A	0.352	0.274	0.548	0.546	0.433	0.142	0.110	0.232	0.227	0.177	0.057	0.045	0.112	0.092	0.072	0.268	0.222	0.292	0.056	0.046
	Prob.H	0.960	0.873	0.999	0.999	1.000	0.076	0.026	0.867	0.873	0.873	0.000	0.000	0.000	0.825	0.833	0.860	0.841	0.894	0.829	0.807
	Mean.H	0.412	0.321	0.812	0.816	0.632	0.148	0.116	0.226	0.313	0.242	0.058	0.045	0.083	0.210	0.163	0.299	0.250	0.328	0.191	0.149
	Med.H	0.412	0.321	0.720	0.771	0.624	0.148	0.116	0.223	0.309	0.242	0.058	0.045	0.083	0.208	0.162	0.299	0.250	0.304	0.189	0.149

Prob.A, Mean.A, Med.A denote coverage frequencies, mean lengths and median lengths of the nominal 90% confidence intervals with the asymptotic expression for standard errors with estimated parameters. Prob.H, Mean.H, Med.H denote coverage frequencies, mean lengths and median lengths of the nominal 90% confidence intervals with the finite sample Hessian based approximation of the standard errors with estimated parameters.

Figure 1: Estimates and CI(95%) of the memory parameter of volatility

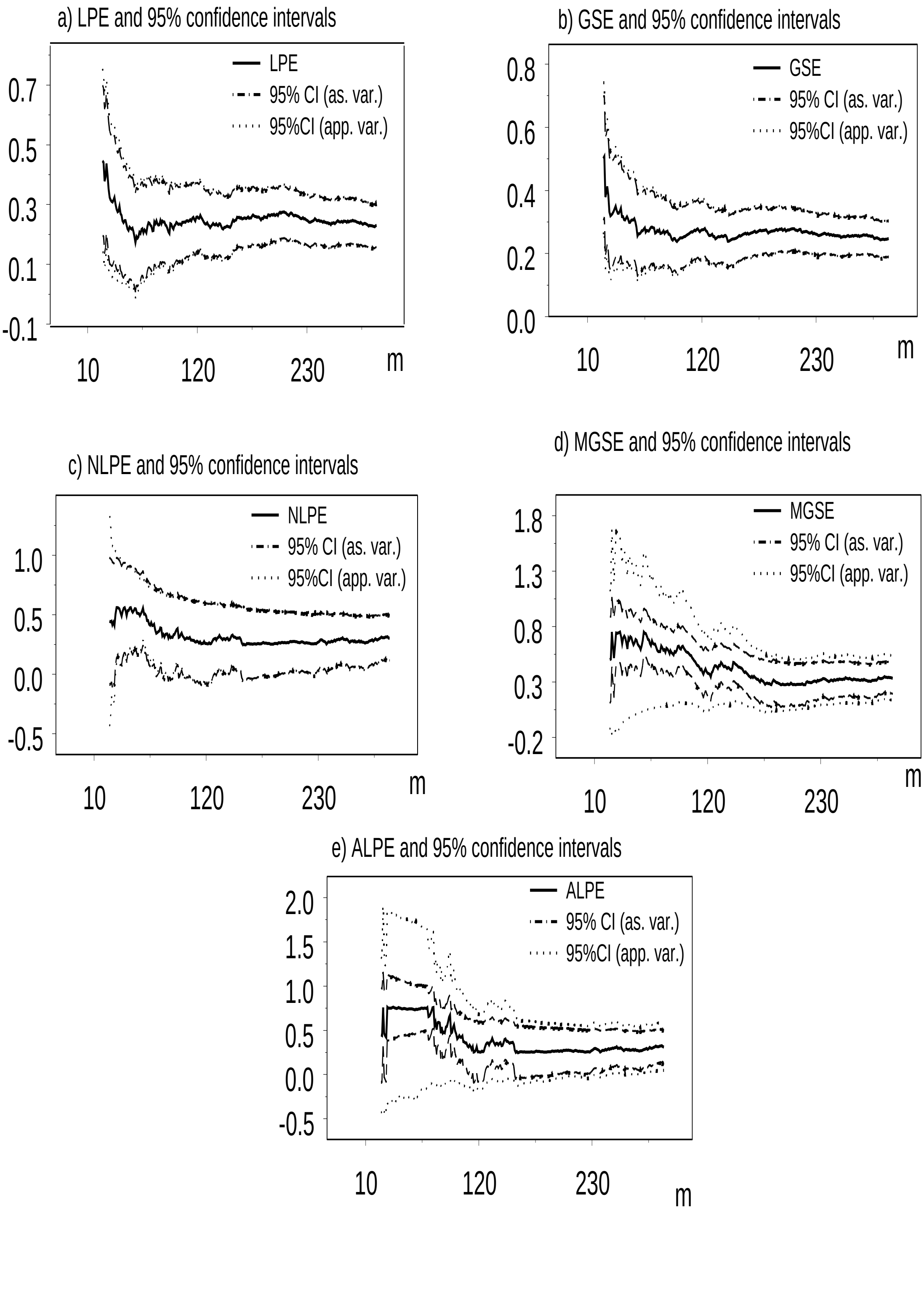

Figure 2: Estimates of the memory parameter of volatility (m=25...200)

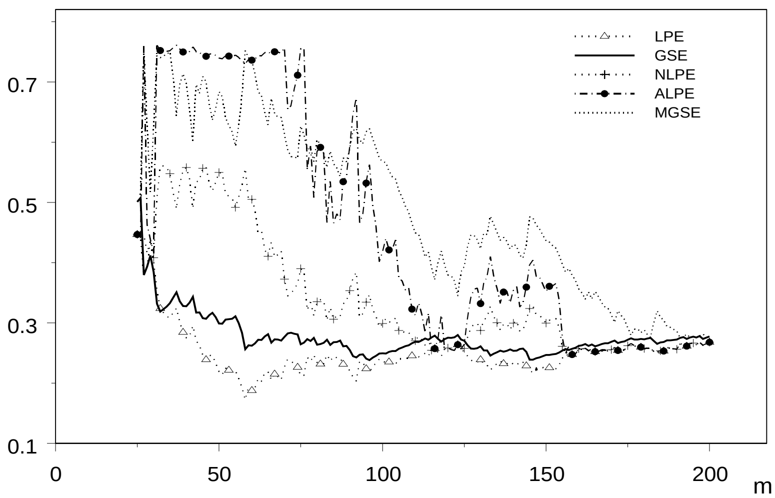

Figure 3: Estimates and CI(95%) of the memory parameter of volatility (m=150...300)