Local Bootstrap Approach for the Estimation of the Memory Parameter

J. Arteche and J. Orbe *

Abstract— **The log periodogram regression is widely used in empirical applications because of its simplicity to estimate the memory parameter, d, its good asymptotic properties and its robustness to misspecification of the short term behavior of the series. However, the asymptotic distribution is a poor approximation of the (unknown) finite sample distribution if the sample size is small. Here the finite sample performance of different nonparametric residual bootstrap procedures is analyzed when applied to construct confidence intervals. In particular, in addition to the basic residual bootstrap the local bootstrap that might adequately replicate the structure that may arise in the errors of the regression is considered when the series shows weak dependence in addition to the long memory component. Bias correcting bootstrap to adjust the bias caused by that structure is also considered.**

Keywords: bootstrap, confidence interval, log periodogram regression, long memory

1 Introduction

Long memory processes are characterized by a strong dependence such that the lag-j autocovariances γ_j decrease hyperbolically as $j \to \infty$

$$\gamma_j \sim Gj^{2d-1}$$

for some finite constant G, d is the memory parameter and $a \sim b$ means that a/b tends in the limit to 1. For $d > 0$, $\sum |\gamma_j| = \infty$ but stationarity is guaranteed as long as $d < 1/2$ and mean reversion holds for $d < 1$. It is also usually assumed that $d > -1/2$, which warrants invertibility. Long memory can alternatively be defined in the frequency domain. A stationary time series process has long memory if its spectral density function $f(\cdot)$ satisfies

$$f(\lambda) \sim C|\lambda|^{-2d} \quad \text{as } \lambda \to 0, \tag{1}$$

for some positive finite constant C. Under positive long memory, which is the most common case in economic and financial series, the spectral density diverges at the origin at a rate governed by d. If $d > 1/2$ the process is not stationary and, by definition, the spectral density does not exist. However pseudo spectral density functions can be similarly defined (e.g. [1]) with a behavior as in (1). One issue of main interest in these processes is the estimation of d. Perhaps the most popular is the log periodogram regression estimator (LPE hereafter) originally proposed by [2] and analyzed in detail in [3] and [4]. The LPE is widely used in empirical applications because of its simplicity, since only a least squares regression is required, its good asymptotic and finite samples properties and its robustness to misspecification of the short term behavior of the series. Taking logarithms of the local specification of the spectral density in (1), the LPE ($\hat{d}$) is obtained by least squares in the regression

$$\log I_j = a + dX_j + u_j, \quad j = 1, ..., m, \tag{2}$$

where $X_j = -2\log\lambda_j$, $a = \log C + c$, $c = 0.577216$ is Euler's constant, $I_j = (2\pi n)^{-1}|\sum_{t=1}^{n} x_t \exp(-it\lambda_j)|^2$ is the periodogram of the series x_t, $t = 1, .., n$, at Fourier frequency $\lambda_j = 2\pi j/n$, n is the sample size, $u_j = \log(I_j f(\lambda_j)^{-1}) - c$ and m represents the bandwidth, that is the number of frequencies used in the estimation. For the asymptotics, this bandwidth has to increase with n but at a slower rate such that the band of frequencies used in the estimation degenerates to zero and the local specification in (1) remains valid. [3] and [4] proved the consistency of $\hat{d}$ in the stationary and invertible region $-0.5 < d < 0.5$, and obtained its limit distribution

$$\sqrt{m}(\hat{d} - d) \xrightarrow{d} N\left(0, \frac{\pi^2}{24}\right). \tag{3}$$

Reference [1] showed that the consistency holds even in the nonstationary region $[0.5, 1)$ and the same limit distribution remains valid for $d \in [0.5, 0.75)$.

In practice the choice of the bandwidth is crucial, a large m decreases the variance at the cost of a higher bias which can be extremely large in some situations, for example in the presence of some short term component such as those analyzed below. The choice of an optimal bandwidth is not a simple task. Some attempts have been made in [5], [6] and [7]. However, the performance of all these procedures is not very satisfactory and the results for a grid of bandwidths are usually shown in empirical applications.

*Authors are with Dpt. Econometría y Estadística, University of the Basque Country, Avda. Lehendakari Agirre 83, 48015 Bilbao, Spain (email: josu.arteche@ehu.es, jesus.orbe@ehu.es). The authors wish to acknowledge financial support from the Spanish *Ministerio de Ciencia y Tecnologa* and *FEDER* grants MTM2006-06550 and SEJ2007-61362/ECON and from the Department of Education of the Basque Government through grant IT-334-07 (UPV/EHU Econometrics Research Group).

The log-periodogram estimation of the memory parameter in economic series raises the problem of the small sample size since many economic time series consist of low frequency, monthly ([8]), quarterly ([9] and [10]) or even yearly ([9] and [10]) data. Furthermore, if the series shows a rich spectral behavior around the origin the bandwidth has to be low enough to avoid a large bias in the estimation of d ([11]). Also the strong seasonality in many quarterly and monthly economic series compels the use of a small bandwidth to avoid distorting influence of neighbouring seasonal spectral poles ([12]). As a result the number of frequencies used in the estimation is small and, as noted in [13], the asymptotic distribution in (3) is a poor approximation of the small sample distribution of $\hat{d}$. In this situation, the bootstrap could be a useful tool to make inference without relying on the asymptotic probability distribution.

In the LPE setup the bootstrap is carried out previously to the definition of the regression model (2) and the bootstrapped dependent variable is then the logarithm of the periodogram of the bootstrap samples. Reference [14] proposed instead to bootstrap directly the residuals in the regression (2) avoiding in that way the necessity to deal with the temporal dependence of x_t with the corresponding computational savings and robustness against misspecification. We focus here in this last approach and analyze the implementation of different bootstraps on the coverage of confidence intervals.

As already mentioned, the LPE can be highly biased in the presence of some weak dependent component. Reference [15] shows that an autoregressive or moving-average component in an ARFIMA model can seriously distort the estimation of the memory parameter with a large bias. The source of the bias is the effect of these short memory components on the spectral behavior around frequency zero such that the approximation (1) is only reliable for frequencies very close to the origin. This weak dependence, not considered in the regression in (2), affects the behavior of u_j such that it shows some remaining structure. To try to capture this structure we consider a version of the local bootstrap of [16], applied also in a similar long memory context by [17], and compare its performance with the nonparametric residual bootstrap. We are also concerned with the effects of the bias in the LPE and evaluate the capacity of different bias corrections usually employed in the bootstrap literature, namely the Bias Corrected (BC) percentile of [18], the accelerated Bias Corrected (BC$_a$) percentile [19] and the Constant Bias Correcting (CBC) estimator of [20]. In addition the bootstrap-t method of [18], which implicitly includes a bias adjustment, is also examined.

The paper is organized as follows. Section II describes the different bootstrap procedures that we analyze in the Monte Carlo in Section III. Finally Section IV concludes.

2 Different bootstraps procedures in LPE

Here the bootstrap to calculate confidence intervals for the LPE that improve standard confidence intervals based on the asymptotic distribution in a small sample size situation is used. We focus on a bootstrap in a regression context, using a nonparametric residual bootstrap since the regressor is based on non stochastic Fourier frequencies and we do not assume any probability distribution for the error term in the regression model, involving a nonparametric resampling of the residuals.

2.1 Residual bootstrap (RB)

The appropriate performance of this bootstrap in the LPE context over other procedures, such as the wild bootstrap, has been discussed in [14]. The *residual bootstrap* for the log periodogram regression follows these steps:

1. Obtain the LPEs, $\hat{a}$, $\hat{d}$, by OLS in the regression (2) and the residuals $\hat{u}_j = \log I_j - \hat{a} - \hat{d}X_j$. Construct the modified residuals $\hat{v}_j = \hat{u}_j/(1-h_j)^{1/2}$.

2. Resampling with replacement from the modified residuals $\hat{v}_j$, and giving equal probability $1/m$ to every residual, get B bootstrap samples $\hat{v}^*_{bj}$, $b = 1, 2, ..., B$ and $j = 1, ..., m$. Using the empirical distribution function of the residuals and based on model (2) we obtain the corresponding bootstrap dependent variable $\log I^*_{bj} = \hat{a} + \hat{d}X_j + \hat{v}^*_{bj}$.

3. Fit the regression model (2) in each bootstrap sample to obtain the B bootstrap estimates $\hat{d}^*_b$, $b = 1, ..., B$.

2.2 Residual local bootstrap (RLB)

The RB implicitly assumes that the errors do not have any structure and their behavior approximate an iid sequence. This can be quite unrealistic, especially when the long memory series contains also a short memory component. We propose here a version of the local bootstrap ([16]) that tries to capture the structure of the errors by bootstrapping only in a neighborhood of each observation. It follows these steps:

1. Step 1 in the RB.

2. Select a resampling width $k_m \in \mathcal{N}$, $k_m \leq [m/2]$ for $[\cdot]$ denoting "the integer part of".

3. Define i.i.d. discrete random variables $S_1, ..., S_m$ taking values in the set $\{0, \pm1, , ..., \pm k_m\}$ with equal probability $1/(2k_m + 1)$.

4. Generate B bootstrap series $\hat{v}^*_{bj} = \hat{v}_{|j+S_j|}$ if $|j+S_j| > 0$, $\hat{v}^*_{bj} = \hat{v}_1$ if $j + S_j = 0$ for $b = 1, 2, ..., B$.

5. Generate B bootstrap samples for the dependent variable $\log I^*_{bj} = \hat{a} + \hat{d}X_j + \hat{v}^*_{bj}$ for $b = 1, 2, ..., B$.

6. Fit the regression model (2) in each bootstrap sample to obtain the B bootstrap estimates $\hat{d}^*_b$, $b = 1, ..., B$.

These bootstrap techniques are used to construct confidence intervals trying to improve the coverage of confidence intervals based on the asymptotic distribution:

$$CI_{(1-\alpha)} = \left(\hat{d} - z_{1-\frac{\alpha}{2}}\hat{se}(\hat{d}); \quad \hat{d} - z_{\frac{\alpha}{2}}\hat{se}(\hat{d})\right)$$

where $\hat{se}(\hat{d})$ is the OLS estimate of the standard error and z_α indicate de $100 \cdot \alpha$th percentile of a $N(0,1)$ distribution. The use of the OLS standard error $\hat{se}(\hat{d})$ instead of the asymptotic variance in (3) has proved to significantly improve the finite sample coverage probabilities.

For each of the bootstrap resampling strategies we consider five different classes of bootstraps confidence intervals for d parameter: the percentile interval (P), the constant bias correction percentile interval (CBC), the bias corrected interval (BC), the accelerated bias corrected interval (BCa) and the bootstrap-t interval (b-t).

1) The basic percentile method (P), proposed by [21]:

$$CI_{(1-\alpha)} = \left(\hat{d}^*_{\left((B+1)(\frac{\alpha}{2})\right)} \quad ; \quad \hat{d}^*_{\left((B+1)(1-\frac{\alpha}{2})\right)}\right).$$

where the $\hat{d}^*_{(j)}$ denotes the jth ordered value of the bootstrap estimates of d.

2) Reference [20] proposed a method for reducing the finite sample bias of consistent estimators using a pre-bootstrap estimation of the bias. The constant bias correcting (CBC) estimator is obtained as $\tilde{d} = \hat{d} - \hat{b}$ where $\hat{b}$ is a bootstrap estimate of the finite sample bias of $\hat{d}$.

$$CI_{(1-\alpha)} = \left(\tilde{d}^*_{\left((B+1)(\frac{\alpha}{2})\right)} \quad ; \quad \tilde{d}^*_{\left((B+1)(1-\frac{\alpha}{2})\right)}\right).$$

3) In order to improve the coverage probability of the basic percentile interval [18] introduced the bias-corrected (BC) percentile. The CI is then constructed as

$$CI_{(1-\alpha)} = \left(\hat{d}^*_{\left((B+1)(\frac{\hat{\alpha}}{2})\right)} \quad ; \quad \hat{d}^*_{\left((B+1)(1-\frac{\hat{\alpha}}{2})\right)}\right),$$

where

$$\frac{\hat{\alpha}}{2} = \Phi\left(2k_0 + z_{\frac{\alpha}{2}}\right) \quad \text{and} \quad 1 - \frac{\hat{\alpha}}{2} = \Phi\left(2k_0 + z_{1-\frac{\alpha}{2}}\right),$$

Φ is the standard normal cumulative distribution function and k_0 is the bias-correction parameter.

4) The accelerated bias-corrected (BC_a) percentile method of [19] is defined as

$$CI_{(1-\alpha)} = \left(\hat{d}^*_{\left((B+1)(\frac{\tilde{\alpha}}{2})\right)} \quad ; \quad \hat{d}^*_{\left((B+1)(1-\frac{\tilde{\alpha}}{2})\right)}\right),$$

where

$$\frac{\tilde{\alpha}}{2} = \Phi\left(k_0 + \frac{k_0 + z_{\frac{\alpha}{2}}}{1 - s\left(k_0 + z_{\frac{\alpha}{2}}\right)}\right)$$

and

$$1 - \frac{\tilde{\alpha}}{2} = \Phi\left(k_0 + \frac{k_0 + z_{1-\frac{\alpha}{2}}}{1 - s\left(k_0 + z_{1-\frac{\alpha}{2}}\right)}\right).$$

If $s = 0$, this confidence interval will be equal to the BC confidence interval. In addition, if the $k_0 = 0$, we are in the basic percentile case.

5) The percentile-t or bootstrap-t method ([18]) is based on a given studentized pivot, in this case: $t = \frac{\hat{d}-d}{\hat{se}(\hat{d})}$. The resulting $(1-\alpha)$ confidence interval is

$$CI = \left(\hat{d} - \hat{se}(\hat{d})t^*_{\left((B+1)(1-\frac{\alpha}{2})\right)}; \hat{d} - \hat{se}(\hat{d})t^*_{\left((B+1)(\frac{\alpha}{2})\right)}\right)$$

For a more detailed description of these and others bootstrap resample procedures and confidence intervals see, for example, [22] or [23].

3 Monte Carlo simulation study

Table 1: LPE 95% confidence intervals coverage for $m = 5$

	AR(1)	p=0.9		AR(1)	p=0.3	
	d=0	d=0.4	d=0.8	d=0	d=0.4	d=0.8
Asym						
	72.8	72.8	76.9	85.9	85.5	86
	1.790	1.756	1.560	1.766	1.761	1.704
RB						
P	70.5	71.4	75.2	84.7	84.9	85.4
	1.734	1.702	1.520	1.714	1.713	1.657
CBC	71	72.3	75.6	84.7	85.3	85.8
	1.769	1.737	1.553	1.749	1.749	1.690
BC	70.9	70.7	75.1	83.9	84.9	85.3
	1.725	1.693	1.511	1.704	1.706	1.648
BCa	70.9	71.4	75.3	84.3	84.7	85.4
	1.740	1.706	1.525	1.719	1.717	1.662
b-t	91.8	88.5	90	95.2	94.8	95.3
	2.849	2.802	2.481	2.813	2.806	2.709
RLB(2)						
P	66	68.6	73.6	90.5	90.1	91.3
	1.668	1.636	1.443	1.631	1.631	1.579
CBC	70.6	70.7	72.4	89.2	89.9	89.9
	1.813	1.781	1.574	1.783	1.789	1.727
BC	62.5	64	64.2	80	80.6	79.3
	1.561	1.524	1.355	1.536	1.518	1.463
BCa	61.5	63.4	64.4	80.8	80	79.3
	1.559	1.527	1.356	1.537	1.518	1.465
b-t	89.4	87.3	89.1	95.5	94.4	94.6
	2.714	2.665	2.383	2.698	2.668	2.574

RB and RLB(k_m) denote the residual bootstrap and the residual local bootstrap with resampling width k_m. In each cell the first number is coverage frequency in percentages over 1000 simulations and the number below it is the average length of the interval. Asym, P, CBC, BC, BCa and b-t denote the confidence intervals based on the asymptotic distribution, the basic percentile, the constant bias correcting percentile, the bias corrected, the accelerated bias corrected and the bootstrap-t respectively.

The performance of the bootstrap in LPE based confidence intervals is assessed in two different type of models:

- **Model 1**: $(1 - 0.9L)(1 - L)^d x_t = \varepsilon_{1t}$

Table 2: LPE 95% confidence intervals coverage for $m = 10$

	AR(1)	p=0.9		AR(1)	p=0.3	
	d=0	d=0.4	d=0.8	d=0	d=0.4	d=0.8
Asym						
	34.7	38.5	47	92.2	89.6	91
	1.097	1.068	0.907	1.085	1.097	1.053
RB						
P	37.7	40.5	49	92.4	90.1	90.6
	1.102	1.073	0.916	1.086	1.100	1.052
CBC	37.1	40.4	48.9	92.5	90.5	91
	1.104	1.074	0.917	1.088	1.102	1.054
BC	39.1	43.5	51.5	92.5	90.5	91.3
	1.102	1.073	0.916	1.087	1.101	1.053
BCa	39	43.4	51.3	92.5	90.4	91.3
	1.102	1.073	0.915	1.086	1.101	1.053
b-t	50	53.6	62.6	95.2	94.4	95.2
	1.297	1.263	1.073	1.285	1.300	1.246
RLB(2)						
P	25.8	28.8	36.7	88.9	85.6	87
	0.954	0.929	0.792	0.929	0.948	0.895
CBC	26.2	30.7	38	87.6	82.9	84.8
	0.967	0.940	0.803	0.942	0.960	0.906
BC	26.1	30.6	38.1	83.7	79.8	81.7
	0.947	0.920	0.793	0.924	0.942	0.892
BCa	26	30.6	38.1	83.7	79.8	81.6
	0.947	0.920	0.794	0.923	0.941	0.892
b-t	44.5	48.3	57.2	93.4	92.7	93.6
	1.275	1.229	1.086	1.236	1.261	1.190
RLB(4)						
P	28.8	32.8	42	93.6	91.8	93
	1.044	1.018	0.862	1.018	1.040	0.990
CBC	33.7	37.3	44.3	90.1	87.2	88.6
	1.053	1.025	0.867	1.032	1.051	1.000
BC	34.3	38.4	44.1	86.1	82.6	86.1
	1.030	1.006	0.856	1.005	1.027	0.974
BCa	34.2	38.1	44	86.1	82.5	86
	1.030	1.006	0.856	1.004	1.027	0.974
b-t	47	52.8	64	97	96	97.2
	1.314	1.278	1.118	1.284	1.310	1.249

- **Model 2**: $(1-0.3L)(1-L)^d x_t = \varepsilon_{1t}$

where L is the lag operator ($Lx_t = x_{t-1}$), ε_{1t} and ε_{2t} are independent standard normal series and $d \in (0, 0.4, 0.8)$. For $d = 0$ the series are short memory such that the spectral density function is positive, bounded and continuos at every frequency. The value $d = 0.4$ corresponds to a stationary long memory series with a spectral density diverging at the origin. For $d = 0.8$ the series is nonstationary and mean reverting. Note that in this case the asymptotic distribution in (3) does not apply and the LPE, although consistent, has a nonnormal limit distribution that depends on d (non pivotal).

These models belong to the ARFIMA class and have a spectral (pseudospectral in the nonstationary case) density function

$$f(\lambda) = \frac{1}{2\pi}\frac{[2\sin\left(\frac{\lambda}{2}\right)]^{-2d}}{|1-\phi e^{-i\lambda}|^2} \sim \frac{1}{2\pi(1-\phi)^2}|\lambda|^{-2d} \quad \text{as } \lambda \to 0$$

for $\phi = 0.9, 0.3$ in Models 1 and 2 respectively. Both include an $AR(1)$ short memory component with moderate (Model 2) and high (Model 1) dependence that gives rise to a bias in the LPE if a large bandwidth is used, especially in Model 1.

Since the bootstrap is essentially beneficial with a low sample size, we only consider $n = 128$, which is comparable to the number of observations in many economic series. For each model three bandwidths are considered $m = 5$, 10 and 20. For the local bootstrap we use different resampling widths $k_m = 2$ (for $m = 5$), $k_m = 2, 4$ (for $m = 10$) and $k_m = 2, 4, 8$ (for $m = 20$). Since the results are very sensitive to the choice of the bandwidth we also consider the plug-in optimal bandwidth proposed by [7] and defined as $m^* = \hat{C}n^{4/5}$ for

Table 3: LPE 95% confidence intervals coverage for $m = 20$

	AR(1)	p=0.9		AR(1)	p=0.3	
	d=0	d=0.4	d=0.8	d=0	d=0.4	d=0.8
Asym						
	2.8	4	20.1	90.8	89.6	87.7
	0.696	0.678	0.522	0.687	0.690	0.676
RB						
P	3.6	5.6	22.9	92.7	91.1	89.4
	0.703	0.685	0.531	0.691	0.696	0.681
CBC	3.7	5.8	22.8	92.4	90.8	89.4
	0.703	0.685	0.531	0.691	0.696	0.681
BC	4.7	7.4	25.2	92.6	91	90.4
	0.705	0.687	0.532	0.693	0.698	0.684
BCa	4.7	7.4	25.2	92.6	91	90.4
	0.705	0.687	0.532	0.693	0.698	0.684
b-t	5.1	7.6	25.6	94.5	92.5	92
	0.749	0.729	0.561	0.738	0.742	0.727
RLB(2)						
P	4	6	22.2	83.4	81	79
	0.625	0.610	0.492	0.594	0.598	0.576
CBC	4.2	6.4	22.6	83.3	80.8	77.7
	0.629	0.614	0.496	0.598	0.602	0.578
BC	4.5	6.6	23.3	81	78.6	76
	0.625	0.613	0.496	0.597	0.601	0.579
BCa	4.5	6.6	23.3	81	78.6	76
	0.625	0.613	0.496	0.597	0.601	0.579
b-t	5.7	7.7	27.5	90.1	87.9	84.7
	0.712	0.698	0.584	0.672	0.681	0.653
RLB(4)						
P	5.2	6.8	25.4	87.5	84.7	83.2
	0.677	0.656	0.531	0.644	0.650	0.626
CBC	5.6	7.7	24.2	87.2	85	82.1
	0.677	0.659	0.532	0.646	0.651	0.627
BC	7.3	9.2	26.2	85.1	82.2	79.3
	0.676	0.657	0.532	0.645	0.650	0.625
BCa	7.3	9.2	26.2	85.1	82.2	79.3
	0.676	0.657	0.532	0.645	0.650	0.625
b-t	6.2	9.1	31.1	92.9	92.1	89.5
	0.764	0.746	0.621	0.723	0.733	0.705
RLB(8)						
P	4.2	6	26.4	90.1	88.8	86.5
	0.698	0.676	0.543	0.677	0.681	0.659
CBC	6.2	8.1	25.7	88.8	86.4	84.4
	0.695	0.672	0.541	0.675	0.679	0.660
BC	10.1	11.4	27.5	86.3	84.8	81.5
	0.700	0.680	0.542	0.679	0.683	0.659
BCa	10.1	11.4	27.5	86.3	84.8	81.5
	0.700	0.680	0.542	0.679	0.683	0.659
b-t	4.7	7.6	32	94.2	92.9	92.8
	0.763	0.747	0.611	0.739	0.746	0.724

$$\hat{C} = \left(\frac{27}{128\pi^2}\right)^{1/5} \hat{K}^{-2/5}$$

where $\hat{K}$ is obtained as the third coefficient in an ordinary linear regression of $\log I_j$ on $(1, -2\log\lambda_j, \lambda_j^2/2)$ for $j = 1, 2, ..., An^{\delta}$, with $4/5 < \delta < 1$ and A an arbitrary constant. Following [7], we use $\delta = 6/7$ and $A = 0.25$. Note that this optimal bandwidth is only consistent for Models 1 and 2 in the stationary region, but we use it also in the rest of cases for illustrative purposes. In practice m^* is obtained as the median of the optimal bandwidths in 1000 series generated in each model. The use of the median instead of the mean avoids the distorting effect of extreme cases. We get in this way $m^* = 12$ and 13 for Models 1 and 2 respectively. The optimal bandwidth is quite robust to different values of d (for large d the optimal bandwidth differs at most one unity from the corresponding optimal bandwidth for low d) and we use the same m^* for all d. The number of bootstraps is $B = 999$ which is large enough for the calculous of confidence intervals ([22]). The number of simulations is 1000.

Tables 1-4 show the coverage frequencies in percentage (first number in each cell) and the average length of the interval (under the frequencies) of confidence intervals for

Table 4: LPE 95% confidence intervals coverage for optimal bandwidth m^*

	AR(1)	p=0.9		AR(1)	p=0.3	
	d=0	d=0.4	d=0.8	d=0	d=0.4	d=0.8
Asym						
	22.6	25.4	39.6	91.5	92.4	89.9
	0.971	0.946	0.762	0.904	0.908	0.883
RB						
P	26.9	28.6	41.8	91.9	92.4	90.7
	0.977	0.952	0.773	0.908	0.914	0.887
CBC	27.1	28.6	41.5	92.1	93	91
	0.978	0.952	0.774	0.909	0.914	0.887
BC	29.3	31.1	43.6	91.8	92.7	90.9
	0.979	0.952	0.774	0.911	0.916	0.889
BCa	29.2	30.9	43.6	91.9	92.6	91
	0.978	0.951	0.774	0.911	0.916	0.888
b-t	35.2	38.6	51.3	94.3	95.3	93.7
	1.109	1.077	0.871	1.018	1.025	0.995
RLB(2)						
P	17.2	20.6	35	86.5	86.7	83.2
	0.840	0.826	0.675	0.785	0.783	0.757
CBC	17.7	21	34.9	83.9	86.2	82
	0.847	0.835	0.683	0.793	0.790	0.765
BC	17.8	21.1	35.1	81.3	82.8	79.8
	0.838	0.826	0.680	0.788	0.787	0.758
BCa	17.8	21.1	35.1	81.3	82.8	79.8
	0.838	0.826	0.680	0.788	0.787	0.758
b-t	31.1	32.9	47.4	92.5	92.8	90.3
	1.059	1.040	0.886	0.975	0.971	0.939
RLB(4)						
P	19.5	23.7	39.8	91.1	91.5	88.3
	0.919	0.906	0.736	0.854	0.856	0.827
CBC	21.8	24.8	38	87.8	89.8	85.9
	0.922	0.910	0.738	0.858	0.862	0.832
BC	23.2	27.7	40.7	83.9	85.2	82.2
	0.910	0.897	0.735	0.850	0.853	0.822
BCa	23.1	27.7	40.7	83.8	85.2	82.1
	0.910	0.897	0.734	0.850	0.853	0.822
b-t	34.4	37.5	53.4	95.2	96	93.6
	1.117	1.101	0.927	1.027	1.030	0.993
RLB(6)						
P	19.9	24.5	40.5	92.7	94.6	91.1
	0.949	0.933	0.755	0.883	0.884	0.859
CBC	25.3	28.1	40.2	88.3	90.6	86.4
	0.954	0.938	0.758	0.888	0.888	0.862
BC	28.1	31	41.9	84.4	86.2	83
	0.938	0.924	0.752	0.878	0.878	0.852
BCa	27.8	31	41.8	84.4	86.2	83
	0.938	0.924	0.752	0.878	0.879	0.852
b-t	31.5	36.7	55.1	96.5	96.8	95
	1.107	1.093	0.910	1.021	1.027	0.993

a 95% nominal confidence level over 1000 Monte Carlo replications with bandwidths $m = 5$, 10, 20 and m^*. The following conclusions can be extracted: i) The bootstrap confidence intervals clearly beats the asymptotic distribution with better coverage frequencies. ii) The different bias correction techniques are only slightly beneficial in Model 1 with a large bandwidth where the bias of the LPE is especially large. The BC and BCa give better results in terms of coverage frequencies than the CBC and the basic P. However the bootstrap-t generally overcomes all the others even in these highly biased situations. iii) The choice of the bandwidth is crucial. The best results are obtained with a low bandwidth when there is a highly dependent short memory component (Model 1) and with a larger bandwidth for models with low dependent short memory component (Model 2). Especially harmful is the use of a large bandwidth ($m = 20$) in Model 1, with very low coverage frequencies. iv) The performance of the local bootstrap depends on the choice of the resampling width k_m. Reference [17] suggested a value of $k_m = 1$ or 2. These values can be too small when the short memory component is of lesser importance and a larger k_m gives better results in these cases. An excessively large k_m can however be harmful in those cases where the estimator is subject to a large bias as in Model 1 with a large bandwidth. Thus a larger k_m should be chosen when the bias component is low. In this situation a large bandwidth should also be used. Then, as a rule of thumb, a larger k_m can be chosen when the optimal bandwidth m^* is large and a low m^* should be accompanied by a small k_m. For the optimal bandwidths in table 4 we found that a value around $k_m = 4$ is adequate. v) The optimal bandwidth of [7] is obtained by minimizing an asymptotic approximation of the mean squared error of the LPE but need not give the best coverage frequencies. This is the case in Models 1 where the bias component is especially large and better coverage frequencies are achieved with a lower bandwidth than the optimal m^* in Table 4. vi) Overall the basic and local residual bootstrap-t give the best performances. Table 5 displays the outcome obtained with the asymptotic distribution and the RB and RLB bootstrap-t with the values of m and the resampling width k_m that give the best coverage frequencies. Note that the optimal bandwidth m^* does not generally correspond to the best performance. The improvements of the bootstrap over the asymptotic distribution are significant. The local bootstrap gives similar coverages to the basic residual bootstrap but with narrower intervals.

Table 5: Best results for coverage frequencies with asymptotic distribution and bootstrap-t

	AR(1)	p=0.9		AR(1)	p=0.3	
	d=0	d=0.4	d=0.8	d=0	d=0.4	d=0.8
Asym						
m	5	5	5	10	m^*	10
cov	72.8	72.8	76.9	92.2	92.4	91
ampl	1.790	1.756	1.560	1.085	0.908	1.053
RB						
m	5	5	5	10	5	10
cov	91.8	88.5	90	95.2	94.8	95.2
ampl	2.849	2.802	2.481	1.285	2.806	1.246
RLB						
m	5	5	5	m^*	5	m^*
k_m	2	2	2	4	2	6
cov	89.4	87.3	89.1	95.2	94.4	95
ampl	2.714	2.665	2.383	1.027	2.668	0.993

4 Conclusion

This paper shows the improvements of some residuals based nonparametric bootstrap strategies over the asymptotic distribution of the LPE in the construction of confidence intervals with a small sample size. It is noteworthy the crucial role played by the choice of the bandwidth. The coverage frequencies and length of the confidence interval vary significantly with m and an appropriate m should be selected as a first step. The RB and the RLB bootstrap-t seems to perform well with an appropriate selection of the resampling width. We have proposed a rule of thumb for approximate selection of the resampling width of the RLB linked to the optimal bandwidth estimation of [7], a high optimal bandwidth requires a high resampling width. The advantage of using the RLB over the RB is the reduction of the length of the confidence intervals without significantly affecting the coverage.

Our analysis has focused on the basic LPE, which is the most popular method of estimation of the memory parameter. There have been recently further refinements either in a linear regression setup or in a nonlinear regression approach. For example [24] proposed a bias reduced LPE by including linearly extra regressors that account for the weak dependent components. This extension can be applied to ARFIMA models such as Models 1 and 2 in our Monte Carlo.

References

[1] Velasco, C., 1999. Non stationary log-periodogram regression. *J. Econometrics* 91, 325-371.

[2] Geweke, J. and Porter-Hudak, S., 1983. The estimation and application of long-memory time series models. *J. Time Ser. Anal.* 4, 221-238.

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

[4] Hurvich, C.M., Deo, R. and 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.

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

[6] Giraitis, L., Robinson, P.M. and Samarov, A., 2000. Adaptive semiparametric estimation of the memory parameter. *J. Multiv. Anal.* 72, 183-207.

[7] Hurvich, C.M., and Deo, R.S., 1999. Plug-in selection of the number of frequencies in regression estimates of the memory parameter of a long-memory time series. *J. Time Ser. Anal.* 20, 331-341.

[8] Hassler, U. and Wolters, J., 1995. Long memory in inflation rates: International evidence. *J. Business Econ. Stat.* 13, 37-45.

[9] Diebold, F.X. and Rudebush, G., 1989. Long memory and persistence in aggregate output. *J. Monet. Econ.* 24, 189-209.

[10] Diebold, F.X. and Rudebush, G., 1991. Is consumption too smooth: Long memory and the Deaton paradox. *Rev. Econ. Statist.* 73, 1-9.

[11] Sowell, F., 1992a. Maximum likelihood estimation of stationary univariate fractionally integrated time series models. *J. Econometrics* 53, 165-188.

[12] Arteche, J. and Robinson, P.M., 2000. Semiparametric inference in seasonal and cyclical long memory processes. *J. Time Ser. Anal.* 21, 1-27.

[13] Sowell, F., 1992b. Modeling long-run behaviour with the fractional ARIMA model. *J. Monet. Econ.* 29, 277-302.

[14] Arteche, J. and Orbe, J., 2005. Bootstrapping the log-periodogram regression. *Econ. Letters* 86, 70-85.

[15] Agiakloglou, C., Newbold, P. and Wohar, M., 1993. Bias in an estimator of the fractional difference parameter. *J. Time Ser. Anal.* 14, 235-246.

[16] Paparoditis, E. and Politis, D., 1999. The local bootstrap for periodogram statistics. *J. Time Ser. Anal.* 20, 193-222.

[17] Silva, E.M., Franco, G.C., Reisen, V.A. and Cruz, F.R.B., 2006. Local bootstrap approaches for fractional differential parameter estimation in ARFIMA models. *Comput. Statist. Data Anal.* 51, 1002-1011.

[18] Efron, B., 1982. The jackknife, the bootstrap, and other resampling plans. Volume 38 of CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM.

[19] Efron, B., 1987. Better bootstrap confidence intervals. *J. Amer. Statistical Assoc.* 82, 171-200.

[20] Mackinnon, J.G. and Smith, A.A., 1998. Approximate bias correction in econometrics. *J. Econometrics* 85, 205-230.

[21] Efron, B., 1979. Bootstrap methods: Another look at the jackknife. *Ann. Statist.* 7, 1-26.

[22] Efron, B. and Tibshirani, R.J., 1993. An introduction to the bootstrap. Chapman and Hall: New York.

[23] Davison, A.C. and Hinkley, D.V., 1997. Bootstrap methods and their application. Cambridge University Press: Cambridge. volatility. *Econometrica* 73, 1283-1328.

[24] Andrews, D.W.K. and Guggenberger, P., 2003. A bias-reduced log-periodogram regression estimator for the long-memory parameter. *Econometrica* 71, 675-712. perturbed fractional processes. *J. Econometrics* 115, 355-389.