The effect of tapering on the semiparametric estimators for nonstationary long memory processes

Mohamed BOUTAHAR Vêlayoudom MARIMOUTOU Leïla NOUIRA *

Université de la Méditerranée

2 rue de la Charité, 13002 Marseille, France.

Abstract

In this paper, we study, by a Monte Carlo simulation, the effect of the order p of "Zhurbenko-Kolmogorov" taper on the asymptotic properties of semiparametric estimators. We show that $p = [d + 1/2] + 1$ gives the smallest variances and mean squared errors. These properties depend also on the truncation parameter m. Moreover, we study the impact of the short-memory components on the bias and variances of these estimators. We finally carry out an empirical application by using four monthly seasonally adjusted logarithm Consumer Price Index series.

Key Words and Phrases: Long-range dependence, Order of tapering, Semiparametric estimators, Monte Carlo study.

AMS 2000 Subject Classification: C13, C14, C15, C22.

1 Introduction

After the seminal papers of Granger and Joyeux (1980) and Hosking (1981), fractional integration processes ($I(d)$) have attracted the attention of many statisticians and econometricians. These long-range dependence processes give more flexibility to empirical research than the classical $I(0)$ and $I(1)$ classes of processes.

For $0 < d < 1/2$, they are stationary with hyperbolic decay of the autocorrelation function and they exhibit long memory or long-range dependence. for $d \geq 1/2$, they are nonstationary. To estimate d, we usually use the semiparametric methods developed by Geweke and Porter-Hudak (1983) (henceforth referred to GPH).[1]

*Correspondance: Leïla NOUIRA, GREQAM, Université de la Méditerranée, Centre de la Vieille Charité, 2 Rue de la Charité, 13002 Marseille, France. Tél.: +33 4 91 14 07 21, fax: +33 4 91 90 02 27. E-mail: leila.nouira@univmed.fr.

[1] They are called semiparametric in the sense that the spectral density is parameterized only within a neighbourhood of zero frequency.

Agiakloglou et al. (1993) showed that this estimator has a large bias. Reisen (1994), Robinson (1994, 1995a, b) and Lobato and Robinson (1996) give some estimators with a small bias.

Many economic time series exhibit a nonstationary behavior, so it will be necessary to extend the concept of long memory to the nonstationary framework, (see Cheung and Lai (1993), Maynard and Phillips (2001) and Phillips (2005)). Hurvich and Ray (1995) argued, by simulation, that the GPH estimator is consistent only when $d < 1$. Kim and Phillips (1999) showed this result theoretically. In the same context, Velasco (1999a) showed the consistency and the asymptotic normality of the Robinson (1995a) estimator for $d \in [1/2, 3/4]$. To overcome the non-consistency problem, Hurvich and Ray (1995) and Velasco (1999a) suggested the use of data tapering, which was first proposed by Cooley and Tukey (1965) and discussed by Cooley et al. (1967) and Jones (1971).[2] This technique has also been used by many authors, such as Hurvich and Chen (2000), Giraitis and Robinson (2003), Sibbertsen (2004), Olhede et al. (2004), among many others.

For any value of d, Velasco (1999a) showed that if the order p of the taper is greater or equal to $[d + 1/2] + 1$, then the estimator is consistent and asymptotically normal. The aim of this paper is to study the effect of p on the properties of semiparametric methods (GPH (1983) and Robinson (1995a, b)) by using a "Zhurbenko-Kolmogorov" taper. We then determine the optimal choice of p. We will study also the impact of the short-memory component on the properties of these estimators and finally, we extend the results given by Boutahar et al. (2006) to the nonstationary case.

The paper is organized as follows. In the next section, we briefly introduce the nonstationary $ARFIMA$ $(0, d, 0)$ process as well as the definition of the tapering procedure. Section 3 describes the GPH and the Robinson (1995a, b) methods. The simulation results are given in section 4. In section 5, we analyze the monthly seasonally adjusted logarithm CPI data for fourth countries, France, Italy, Germany and U.S.. Section 6 concludes the paper.

2 Nonstationary time series and data tapers

Let $\{X_t\}$ be the $ARFIMA$ $(0, d, 0)$ process generated by

$$(1 - B)^d X_t = \varepsilon_t, \qquad \varepsilon_t \sim i.i.d. \left(0, \sigma_\varepsilon^2\right), \ d \in I\!R, \tag{2.1}$$

where

$$(1 - B)^d = \sum_{k=0}^{\infty} \left(\begin{array}{c} d \\ k \end{array} \right) (-B)^k = 1 - dB - \frac{1}{2} d(1 - d) B^2 - \frac{1}{6} d(1 - d)(2 - d) B^3 .... \tag{2.2}$$

For $-1/2 < d < 1/2$, $\{X_t\}$ is stationary and invertible. Its spectral density $f(\lambda)$ satisfies

[2] The tapering was originally used in nonparametric spectral analysis of short-memory time series to reduce the bias. It can be used in both stationary and nonstationary cases.

$$f(\lambda) \sim G|\lambda|^{-2d} \qquad as \ \lambda \to 0^+, \tag{2.3}$$

where G is a positive constant.[3] When $d \geq 1/2$, let $s = \left[d + \frac{1}{2}\right]$, following Hurvich and Ray (1995), a nonstationary process $\{X_t\}$ exhibits a long memory if $\Delta^s X_t = \varepsilon_t^{(s)}$, is stationary with spectral density satisfying

$$f_{\varepsilon^{(s)}}(\lambda) \sim G\lambda^{-2(d-s)} \qquad as \ \lambda \to 0^+. \tag{2.4}$$

In the nonstationary case, $f(\lambda)$ does not exist, but we can define a "pseudo spectral density" as following[4]

$$f(\lambda) = |1 - \exp(i\lambda)|^{-2s} f_{\varepsilon^{(s)}}(\lambda). \tag{2.5}$$

To reduce the bias of the periodogram, Cooley and Tukey (1965) suggested the use of tapering. It consists of multiplying the data by a sequence of non-negative weights, called a "taper" or "fader" or "data window". It takes values around 1 for the central part of the data, and decays smoothly to 0 at both the beginning and the end of the sample. The tapered discrete Fourier transform of $\{X_t\}_{t=1,2,...,n}$, for any taper sequence $\{h_t\}_{t=1}^n$, is defined as

$$w^T(\lambda_{j,n}) = \left(2\pi \sum_{t=1}^n h_t^2\right)^{-1/2} \sum_{t=1}^n h_t X_t \exp(i\lambda_{j,n} t), \tag{2.6}$$

where $\lambda_{j,n} = \frac{2\pi j}{n}$, $\forall \ j = 1, 2, ..., \left[\frac{n-1}{2}\right]$, and then the tapered periodogram is $I^T(\lambda_{j,n}) = \left|w^T(\lambda_{j,n})\right|^2$. The usual discrete Fourier transform $w(\lambda_{j,n})$ is obtained by setting $h_t \equiv 1 \ \forall \ t$. There exist many data tapers, such as, the "cosine bell" taper, the "Zhurbenko-Kolmogorov" taper and the "Parzen" taper (see Brillinger (1975), Alekseev (1996) and Velasco (1999*a*), for more examples). To implement the taper, we must know its order p, (a positive integer related to the smoothness of the taper).

Definition 1. *A sequence of taper $\{h_t\}_{t=1}^n$ is of order p, if the following two conditions are satisfied*

1. $\sum_{t=1}^n h_t^2 = nb(n)$, *where* $0 < b(n) < \infty$,

2. For $N = [n/p]$, *the Dirichlet Kernel* $D^T(\lambda)$ *satisfies*

$$D^T(\lambda) \equiv \sum_{t=1}^n h_t \exp\{i\lambda t\} = \frac{a(\lambda)}{n^{p-1}} \left(\frac{\sin[n\lambda/2p]}{\sin[\lambda/2]}\right)^p, \tag{2.7}$$

[3] A more detailed description of a stationary $ARFIMA(0, d, 0)$ processes can be found in Beran (1994) and Boutahar et al. (2006).

[4] For more details concerning this nonstationary process, see for example Velasco (1999*a*).

where $a(\lambda)$ is a complex function and $[.]$ is the integer part.

In this paper we are only interested by the "Zhurbenko-Kolmogorov" taper $\left\{h_t^{ZK}\right\}$. It is based on p convolutions of the uniform density in $1, ..., n$.[5] This taper is proportional to the coefficients $c_{p,N}(t)$ defined as

$$\sum_{t=0}^{p(N-1)} z^t c_{p,N}(t+1) = \left(1 + z + ... + z^{N-1}\right)^p = \left(\frac{1 - z^N}{1 - z}\right)^p.$$

When $p = 1$, $h_t^{ZK} \equiv 1$, so no tapering is performed. When $p = 2$, the taper sequence is identical to "Bartlett" or "triangular window"

$$h_t^{ZK} = 1 - \left|\frac{2t - n}{n}\right|, \qquad t = 1, 2, ..., n. \tag{2.8}$$

When $p = 3$, the $\left\{h_t^{ZK}\right\}$ is the same as the "full cosine bell" taper

$$h_t^{ZK} = \frac{1}{2}\left\{1 - \cos(2\pi t/n)\right\}, \qquad t = 1, 2, ..., n. \tag{2.9}$$

Finally when $p = 4$, it is very close to the "Parzen" taper

$$h_t^{ZK} = \begin{cases} 1 - 6\left[\left|\frac{2t-n}{n}\right|^2 - \left|\frac{2t-n}{n}\right|^3\right], & N < t < 3N, \\ 2\left\{1 - \left|\frac{2t-n}{n}\right|\right\}^3, & t \leq N \text{ or } 3N \leq t \leq 4N, \end{cases} \tag{2.10}$$

with $N = [n/4]$.[6]

3 The semiparametric methods

In this section, we are focusing on the GPH and Robinson (1995*a*, *b*) methods. We consider a tapered periodogram by using the Fourier frequencies $\lambda_{j,n}$, where j is a multiple of the order p (see Velasco (1999*a*)).

3.1 The Geweke and Porter-Hudak method

Geweke and Porter-Hudak (1983) have proposed an estimator based on log-periodogram regression. Let

$$I^T(\lambda_{j,n}) = \frac{1}{2\pi \sum_{t=1}^n h_t^2}\left|\sum_{t=1}^n h_t X_t \exp(i\lambda_{j,n} t)\right|^2, \tag{3.11}$$

[5] A good illustration is given in Velasco (1999*a*).

[6] For the graphical representations of the "Zhurbenko-Kolmogorov" taper with $p = 1, 2, 3$ and 4, see Velasco (1999*a*), page 345.

be the tapered periodogram of the process X_t at frequency $\lambda_{j,n} = 2\pi j/n$, where $j = p$, $2p, ..., m$.[7] The spectral regression of the tapered GPH estimator is computed by regressing a number of log tapered periodograms on a constant and a nonlinear function of the frequencies,

$$\ln\left\{I^T\left(\lambda_{j,n}\right)\right\} = a - d\ln\left\{4\sin^2\left(\lambda_{j,n}/2\right)\right\} + e_j, \qquad e_j \sim i.i.d.\left(-c, \pi^2/6\right), \tag{3.12}$$

$c = 0.5772...$, is the Euler's constant. The GPH estimator is given by,

$$\hat{d}^p_{GPH} = \frac{\sum_{j=p,2p,..,m}\left(Y_j - \bar{Y}\right)\ln\left\{I^T\left(\lambda_{j,n}\right)\right\}}{\sum_{j=p,2p,..,m}\left(Y_j - \bar{Y}\right)^2}, \tag{3.13}$$

where
$\mathrm{Y}_j = -\ln\left\{4\sin^2\left(\lambda_{j,n}/2\right)\right\}$, and $\bar{Y} = \left(1/\left[m/p\right]\right)\sum_{j=p,2p,..,m} Y_j$.

This estimator is consistent for $-1/2 < d < 1/2$ and is obtained without knowledge on the distribution of the data generating process (DGP), moreover it does not require any knowledge of the short-term component. However, it has some potential problems, such that, its bias and variance depend on the number of frequencies m used for the estimation. Usually, $m = n^{0.5}$ which may not be the best choice and may lead to biased results. There exists some theoretical work on this topic, but there is no easily applicable rule for the choice of m.[8]

3.2 The Robinson (1995a) method

In order to reduce the bias of GPH estimator, Robinson (1995*a*) proposed a modified version of this estimator, which discards the l first frequencies.[9] The tapered Robinson (1995*a*) estimator is given by

$$\hat{d}^p_{Ra} = \frac{\sum_{j=(1+l)p,(2+l)p,..,m}\left(Y_j - \bar{Y}\right)\ln\left\{I\left(\lambda_{j,n}\right)\right\}}{\sum_{j=(1+l)p,(2+l)p,..,m}\left(Y_j - \bar{Y}\right)^2}, \qquad 0 < l < m < n, \tag{3.14}$$

with

$$\bar{Y} = \frac{1}{[m/p] - l}\sum_{j=(1+l)p,2p,..,m} Y_j. \tag{3.15}$$

There is no optimal choice for the parameters l and m, which induces an important problem in the implementation of this method and can increase the bias of the estimator.

[7] m is equal to $g(n)$, where $\lim_{n\to\infty} g(n) = \infty$, $\lim_{n\to\infty} g(n)/n = 0$ for example $g(n) = n^{\alpha}$, with $0 < \alpha < 1$. It is the number of frequencies that must take a value lesser than $n/2$.

[8] Taqqu and Teverovsky (1996) suggested plotting the estimates of d as a function of m which balances bias versus variance and, if the plot flattens in a central region in which both the variance and the bias of the estimate of d should be small, then we use the flat part for estimating d. Hauser (1997) showed that, in the stationary case and in the presence of strong short-run effects, this bias in $\hat{d}$ may be devastating.

[9] In this paper, we use for simulation and application, $l = 4$.

3.3 The Robinson (1995b) method

Let $Q^p(G,d)$ be the objective function

$$Q^p(G,d) = \frac{p}{m} \sum_{j=p,2p,..,m} \left\{ \log\left(G\lambda_{j,n}^{-2d}\right) + \frac{I^T(\lambda_{j,n})}{G\lambda_{j,n}^{-2d}} \right\}, \tag{3.16}$$

and let Δ_1 and Δ_2, the lower and upper bound of the admissible values of d, such that, $-\infty < \Delta_1 < \Delta_2 < +\infty$,[10] then we define the estimates of (G^p,d^p) as following

$$\left(\hat{G}^p, \hat{d}^p_{Rb}\right) = arg \min_{0<G<\infty, d\in[\Delta_1,\Delta_2]} Q^p(G,d).$$

It can be shown that

$$\hat{d}^p_{Rb} = \arg \min_{d\in[\Delta_1,\Delta_2]} R^p(d), \tag{3.17}$$

where

$$R^p(d) = \log \hat{G}^p(d) - 2d\frac{p}{m}\sum_{j=p,2p,..,m} \log \lambda_{j,n}, \text{ and } \hat{G}^p(d) = \frac{p}{m}\sum_{j=p,2p,..,m} \lambda_{j,n}^{2d} I^T(\lambda_{j,n}).$$

This estimator requires the Normality of the process and its bias depends on the value of the truncation parameter m.

3.4 General comments

In the stationary case, Hauser (1997) studied the properties of some semiparametric and nonparametric tests for the fractional integration parameter d. He showed that the trimmed Whittle likelihood exhibits high power for pure fractionaly integrated processes. Recently, Boutahar et al. (2006) compared, by simulation, the performance of the various classes of estimators. They found that only the semiparametric and the maximum likelihood methods yield good results. Their main result is that the Robinson (1995*b*) method is to be prefered, since it gives the smallest bias and mean squared errors.

Recall that Velasco (1999a, b), Kim and Phillips (1999) and Shimotsu and Phillips (2004) studied the properties of these three estimators in the nonstationary case. They showed that the consistency and the asymptotic normality properties can be obtained only for $d < 3/4$. Moreover, Velasco (1999a, b) showed that the use of an adequate data taper, can extend these properties to any values of the long memory parameter d.

Note that, the semiparametric estimators based on the data tapering are robust to the presence of trends and structural breaks. However, they usually have a larger variances (1.5 times or more) than the untapered semiparametric ones.

[10] In Robinson (1995*b*), Δ_1 and Δ_2 can take a values $-1/2$ and 1 respectively.

4 Monte Carlo Study

We generate 10000 realizations of a Gaussian $ARFIMA$ $(0, d, 0)$ process, taking d in the interval $[0.5, 2.0]$ with steps of 0.1.[11] For some values of d $(0.6, 1.2, 1.8)$, we vary the short-memory parameters ϕ and θ $(0.4, 0.8, 0.9, 0.95)$, and we simulate again an $ARFIMA$ $(1, d, 0)$ as well as an $ARFIMA$ $(0, d, 1)$.[12] We use three sample sizes $n = 100$, 500 and 1500, two values of the truncation parameters m ($m = n^{0.5}$ and $m = n^{0.8}$) and four orders $p = 1$, 2, 3 and 4.[13] For each realization, we compute the three estimators, $((3.13), (3.14)$ and $(3.17))$, hence we determine their average values $\left(\hat{d}^p\right)$ (over the 10000 realizations) and their mean squared errors (MSE). The results of the $ARFIMA$ $(0, d, 0)$ process are grouped in tables $1 - 6$,[14] whereas those of the $ARFIMA$ $(1, d, 0)$ and $ARFIMA$ $(0, d, 1)$ processes are in tables $8 - 11$.

4.1 The analysis of the ARFIMA (0,d,0) results

The comparison of the three semiparametric estimators for different values of d and m, shows that the Robinson (1995*b*) estimator gives the smallest bias and mean squared errors. This result coincides with the stationary case, which was pointed out by Hauser (1997) and Boutahar et al. (2006). We conclude that this estimator is the best for both stationary and nonstationary cases.

- The bias and the MSE decrease as the sample size increases.
- For $d \in [0.5, 1.0]$, the order $p = 1$ can give a good result. Velasco (1999*a*) and Kim and Phillips (1999) showed that for this interval no tapering is needed to obtain a consistent estimator for d. This result is confirmed by our simulation. As we explained in section 2, when the order is 1, no tapering is performed, therefore for this interval we do not need a taper to obtain a good results. We note that with $p = 2$, we obtain a smaller bias than with $p = 1$ but a larger MSE.
- For $1.1 \leq d < 1.5$, the order p must be equal to 2, the MSE increases with p and the bias decreases. For example (see table 5), for the Robinson (1995*b*) method, with $d = 1.2$, $p = 2$, $n = 500$ and $m = n^{0.5}$, the bias is 0.033, the standard error is 0.250 and the MSE is 0.064, whereas with $p = 4$, they are respectively -0.005, 0.460 and 0.212. In the same context and for $1.5 \leq d < 2.0$, the order p must be equal to 3. However for $d = 2.0$, tables 1-6 show that we need only a taper of order 2.

[11] In our choice, we are restricted to $d = 2.0$ because, for a higher value of a long memory parameter d, the deterministic component totally dominate the random one.

[12] In the stationary and nonstationary cases, these processes have the same memory parameter than the $ARFIMA$ $(0, d, 0)$ ones.

[13] The series are simulated with the S-Plus function **arima.fracdiff.sim**.

[14] To save space, only the results for sample size $n = 500$ are given here, the others are supplied upon request.

- If $p < [d+1/2]+1$, then the tapered estimators ($\hat{d}^p_{GPH}$, $\hat{d}^p_{Ra}$, $\hat{d}^p_{Rb}$) converge always to p. When $p \geq [d+1/2]+1$, the bias of all estimators are positive (see tables 1 and 2). For example, with $n = 500$, $d = 1.6$ and $m = n^{0.5}$, the bias of GPH estimator is 0.100 for $p = 3$ and 0.080 for $p = 4$, and they are respectively 0.024 and 0.019 for $m = n^{0.8}$.

Table 1 The GPH estimator with "Zhurbenko Kolmogorov" taper: $ARFIMA\ (0, d, 0)$ with $\sigma = 1$.

$m = n^{0.5}$ and $n = 500$												
	$p = 1$			$p = 2$			$p = 3$			$p = 4$		
d	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE
0.5	0.526	0.172	0.030	0.515	0.272	0.074	0.461	0.375	0.141	0.448	0.497	0.247
0.6	0.634	0.173	0.031	0.620	0.273	0.075	0.563	0.376	0.142	0.551	0.497	0.247
0.7	0.742	0.173	0.031	0.724	0.270	0.073	0.664	0.375	0.142	0.656	0.496	0.246
0.8	0.841	0.173	0.031	0.834	0.267	0.073	0.770	0.375	0.142	0.760	0.496	0.246
0.9	0.930	0.167	0.028	0.942	0.265	0.072	0.939	0.376	0.143	0.865	0.498	0.248
1.0	1.006	0.159	0.025	1.052	0.263	0.072	1.045	0.378	0.145	0.971	0.499	0.249
1.1	1.047	0.149	0.025	1.162	0.263	0.073	1.151	0.384	0.150	1.133	0.501	0.252
1.2	1.070	0.147	0.038	1.274	0.263	0.075	1.258	0.386	0.153	1.240	0.500	0.251
1.3	1.074	0.152	0.074	1.386	0.267	0.080	1.367	0.385	0.153	1.348	0.501	0.253
1.4	1.060	0.143	0.136	1.500	0.273	0.085	1.476	0.383	0.153	1.456	0.510	0.263
1.5	1.057	0.137	0.215	1.613	0.276	0.089	1.587	0.382	0.153	1.569	0.505	0.260
1.6	1.072	0.146	0.301	1.728	0.281	0.095	1.700	0.378	0.153	1.680	0.496	0.252
1.7	1.078	0.156	0.411	1.839	0.283	0.099	1.812	0.377	0.155	1.792	0.492	0.250
1.8	1.074	0.153	0.551	1.946	0.267	0.093	1.925	0.379	0.159	1.905	0.490	0.251
1.9	1.069	0.143	0.712	2.037	0.258	0.085	2.040	0.385	0.167	2.018	0.488	0.252
2.0	1.065	0.147	0.896	2.109	0.246	0.073	2.153	0.386	0.173	2.131	0.489	0.256

Table 2 The GPH estimator with "Zhurbenko Kolmogorov" taper: $ARFIMA\ (0, d, 0)$ with $\sigma = 1$.

$m = n^{0.8}$ and $n = 500$												
	$p = 1$			$p = 2$			$p = 3$			$p = 4$		
d	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE
0.5	0.512	0.060	0.004	0.504	0.090	0.008	0.506	0.111	0.012	0.508	0.130	0.017
0.6	0.617	0.062	0.004	0.605	0.089	0.008	0.607	0.112	0.012	0.608	0.130	0.017
0.7	0.723	0.065	0.005	0.707	0.089	0.008	0.708	0.112	0.012	0.709	0.130	0.017
0.8	0.829	0.068	0.005	0.807	0.088	0.008	0.809	0.112	0.013	0.809	0.130	0.017
0.9	0.926	0.063	0.005	0.909	0.088	0.008	0.910	0.112	0.013	0.907	0.130	0.017
1.0	1.002	0.054	0.003	1.012	0.088	0.008	1.012	0.112	0.013	1.010	0.131	0.017
1.1	1.040	0.060	0.007	1.115	0.088	0.008	1.113	0.113	0.013	1.108	0.131	0.017
1.2	1.051	0.078	0.028	1.218	0.088	0.008	1.215	0.114	0.013	1.210	0.131	0.017
1.3	1.045	0.089	0.073	1.322	0.088	0.008	1.317	0.114	0.013	1.312	0.131	0.017
1.4	1.033	0.089	0.143	1.426	0.090	0.009	1.419	0.113	0.013	1.414	0.133	0.018
1.5	1.025	0.075	0.231	1.531	0.091	0.009	1.521	0.113	0.013	1.516	0.132	0.018
1.6	1.029	0.088	0.333	1.639	0.094	0.010	1.624	0.113	0.013	1.619	0.131	0.017
1.7	1.031	0.090	0.456	1.747	0.097	0.012	1.727	0.113	0.014	1.722	0.130	0.017
1.8	1.026	0.087	0.607	1.857	0.097	0.013	1.830	0.113	0.014	1.824	0.130	0.017
1.9	1.021	0.067	0.777	1.957	0.093	0.012	1.933	0.113	0.014	1.927	0.129	0.018
2.0	1.020	0.071	0.965	2.031	0.086	0.008	2.037	0.113	0.014	2.030	0.129	0.018

Table 3 The Robinson $(1995a)$ estimator with "Zhurbenko Kolmogorov" taper: $ARFIMA\ (0, d, 0)$ with $\sigma = 1$.

$m = n^{0.5}$, $l = 4$ and $n = 500$												
	$p = 1$			$p = 2$			$p = 3$			$p = 4$		
d	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE
0.5	0.521	0.302	0.092	0.507	0.728	0.530	0.516	1.299	1.688	0.460	2.003	4.014
0.6	0.629	0.304	0.093	0.608	0.727	0.529	0.620	1.298	1.685	0.565	2.004	4.017
0.7	0.735	0.301	0.092	0.709	0.729	0.532	0.722	1.300	1.690	0.665	2.005	4.021
0.8	0.836	0.304	0.094	0.811	0.730	0.533	0.827	1.301	1.693	0.770	2.006	4.025
0.9	0.928	0.305	0.094	0.913	0.731	0.535	0.863	1.299	1.689	0.876	2.007	4.029
1.0	0.996	0.302	0.091	1.016	0.731	0.535	0.964	1.299	1.689	0.978	2.008	4.028
1.1	1.034	0.296	0.092	1.118	0.729	0.532	1.064	1.299	1.689	1.077	2.009	4.037
1.2	1.051	0.260	0.090	1.224	0.728	0.531	1.165	1.300	1.691	1.176	2.011	4.045
1.3	1.050	0.235	0.118	1.332	0.728	0.531	1.264	1.301	1.694	1.278	2.013	4.053
1.4	1.036	0.185	0.167	1.440	0.729	0.533	1.361	1.302	1.697	1.379	2.015	4.060
1.5	1.035	0.168	0.244	1.544	0.731	0.536	1.458	1.303	1.700	1.478	2.018	4.073
1.6	1.036	0.155	0.342	1.655	0.733	0.540	1.556	1.303	1.700	1.677	2.020	4.086
1.7	1.033	0.152	0.468	1.775	0.735	0.536	1.654	1.303	1.700	1.788	2.021	4.092
1.8	1.031	0.160	0.618	1.871	0.737	0.548	1.751	1.302	1.698	1.900	2.017	4.078
1.9	1.021	0.130	0.790	1.954	0.741	0.552	1.849	1.303	1.700	2.010	2.010	4.052
2.0	1.019	0.117	0.976	2.012	0.893	0.798	1.963	1.412	1.995	2.016	2.030	4.121

Table 4 The Robinson $(1995a)$ estimator with "Zhurbenko Kolmogorov" taper: $ARFIMA\ (0, d, 0)$ with $\sigma = 1$.

$m = n^{0.8}$, $l = 4$ and $n = 500$												
	$p = 1$			$p = 2$			$p = 3$			$p = 4$		
d	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE	$\hat{d}^p_{GPH}$	$S\hat{E}_d$	MSE
0.5	0.510	0.074	0.006	0.497	0.124	0.015	0.496	0.177	0.031	0.494	0.221	0.049
0.6	0.615	0.075	0.006	0.597	0.124	0.015	0.596	0.177	0.031	0.594	0.221	0.049
0.7	0.721	0.077	0.006	0.697	0.124	0.015	0.696	0.177	0.032	0.694	0.221	0.049
0.8	0.827	0.079	0.007	0.796	0.124	0.015	0.805	0.178	0.032	0.794	0.221	0.049
0.9	0.926	0.075	0.006	0.895	0.124	0.015	0.905	0.178	0.032	0.894	0.221	0.049
1.0	1.001	0.066	0.004	0.995	0.124	0.015	1.006	0.178	0.032	0.994	0.221	0.049
1.1	1.036	0.069	0.009	1.110	0.124	0.015	1.108	0.178	0.032	1.094	0.221	0.049
1.2	1.046	0.083	0.031	1.213	0.123	0.015	1.208	0.178	0.032	1.194	0.221	0.049
1.3	1.037	0.089	0.077	1.317	0.124	0.015	1.307	0.178	0.032	1.294	0.221	0.049
1.4	1.025	0.087	0.148	1.422	0.124	0.015	1.410	0.178	0.032	1.394	0.222	0.049
1.5	1.017	0.071	0.239	1.528	0.126	0.016	1.512	0.178	0.032	1.494	0.223	0.050
1.6	1.018	0.082	0.346	1.635	0.128	0.016	1.613	0.178	0.032	1.595	0.222	0.049
1.7	1.017	0.080	0.473	1.742	0.130	0.017	1.720	0.178	0.032	1.690	0.221	0.049
1.8	1.012	0.078	0.626	1.848	0.132	0.018	1.820	0.178	0.032	1.790	0.221	0.049
1.9	1.007	0.055	0.800	1.946	0.133	0.019	1.922	0.178	0.032	1.890	0.220	0.049
2.0	1.007	0.059	0.989	2.005	0.127	0.016	2.001	0.178	0.032	1.995	0.220	0.048

Table 5 The Robinson $(1995b)$ estimator with "Zhurbenko Kolmogorov" taper: $ARFIMA\ (0, d, 0)$ with $\sigma = 1$.

$m = n^{0.5}$ and $n = 500$												
	$p=1$			$p=2$			$p=3$			$p=4$		
d	$\hat{d}^p_{GPH}$	$\hat{SE}_d$	MSE	$\hat{d}^p_{GPH}$	$\hat{SE}_d$	MSE	$\hat{d}^p_{GPH}$	$\hat{SE}_d$	MSE	$\hat{d}^p_{GPH}$	$\hat{SE}_d$	MSE
0.5	0.507	0.145	0.021	0.484	0.253	0.064	0.488	0.342	0.118	0.485	0.457	0.212
0.6	0.612	0.146	0.021	0.582	0.252	0.064	0.590	0.343	0.119	0.590	0.457	0.212
0.7	0.717	0.145	0.021	0.688	0.251	0.063	0.690	0.343	0.119	0.690	0.458	0.211
0.8	0.822	0.142	0.021	0.795	0.251	0.063	0.790	0.344	0.119	0.814	0.458	0.212
0.9	0.919	0.136	0.019	0.906	0.250	0.063	0.890	0.344	0.119	0.920	0.461	0.214
1.0	0.989	0.127	0.016	1.012	0.250	0.063	0.990	0.345	0.119	1.026	0.461	0.213
1.1	1.047	0.120	0.017	1.122	0.250	0.063	1.090	0.347	0.120	1.082	0.459	0.211
1.2	1.080	0.121	0.029	1.233	0.250	0.064	1.210	0.348	0.121	1.195	0.460	0.212
1.3	1.089	0.129	0.061	1.335	0.252	0.065	1.285	0.350	0.123	1.291	0.463	0.215
1.4	1.077	0.130	0.121	1.446	0.255	0.067	1.419	0.350	0.123	1.394	0.472	0.222
1.5	1.075	0.133	0.199	1.551	0.256	0.070	1.529	0.348	0.122	1.514	0.468	0.219
1.6	1.089	0.144	0.282	1.651	0.255	0.072	1.640	0.346	0.122	1.618	0.464	0.216
1.7	1.098	0.157	0.387	1.765	0.250	0.072	1.748	0.345	0.122	1.720	0.463	0.215
1.8	1.095	0.153	0.521	1.862	0.242	0.069	1.855	0.344	0.123	1.819	0.462	0.214
1.9	1.090	0.151	0.679	1.950	0.232	0.064	1.981	0.344	0.125	1.918	0.461	0.213
2.0	1.084	0.159	0.863	2.081	0.222	0.056	2.064	0.461	0.217	2.095	0.345	0.128

Table 6 The Robinson $(1995b)$ estimator with "Zhurbenko Kolmogorov" taper: $ARFIMA\ (0, d, 0)$ with $\sigma = 1$.

$m = n^{0.8}$ and $n = 500$												
	$p=1$			$p=2$			$p=3$			$p=4$		
d	$\hat{d}^p_{Rb}$	$\hat{SE}_d$	MSE	$\hat{d}^p_{Rb}$	$\hat{SE}_d$	MSE	$\hat{d}^p_{Rb}$	$\hat{SE}_d$	MSE	$\hat{d}^p_{Rb}$	$\hat{SE}_d$	MSE
0.5	0.505	0.050	0.003	0.497	0.073	0.005	0.502	0.092	0.008	0.495	0.110	0.012
0.6	0.610	0.051	0.003	0.600	0.073	0.005	0.604	0.092	0.008	0.596	0.110	0.012
0.7	0.715	0.053	0.003	0.702	0.073	0.005	0.702	0.092	0.008	0.695	0.110	0.012
0.8	0.817	0.054	0.003	0.803	0.073	0.005	0.797	0.092	0.008	0.796	0.110	0.012
0.9	0.914	0.050	0.003	0.903	0.073	0.005	0.897	0.092	0.008	0.899	0.110	0.012
1.0	1.000	0.042	0.002	1.007	0.073	0.005	0.999	0.092	0.008	1.000	0.110	0.012
1.1	1.048	0.047	0.005	1.099	0.073	0.005	1.099	0.092	0.009	1.102	0.110	0.012
1.2	1.068	0.067	0.022	1.205	0.074	0.006	1.203	0.092	0.009	1.204	0.110	0.012
1.3	1.065	0.084	0.062	1.306	0.074	0.006	1.303	0.092	0.009	1.305	0.110	0.012
1.4	1.050	0.090	0.130	1.408	0.075	0.006	1.406	0.092	0.009	1.407	0.110	0.012
1.5	1.041	0.084	0.218	1.513	0.076	0.007	1.509	0.092	0.009	1.509	0.110	0.012
1.6	1.044	0.096	0.318	1.615	0.078	0.007	1.606	0.092	0.009	1.606	0.110	0.012
1.7	1.046	0.102	0.438	1.723	0.080	0.008	1.705	0.092	0.009	1.706	0.110	0.012
1.8	1.040	0.097	0.586	1.828	0.080	0.009	1.806	0.092	0.009	1.807	0.110	0.012
1.9	1.035	0.083	0.755	1.931	0.076	0.008	1.904	0.093	0.009	1.908	0.110	0.013
2.0	1.032	0.089	0.944	2.025	0.069	0.005	2.034	0.093	0.010	2.023	0.111	0.013

- The order of taper has also an impact on the distribution of these three estimators. To verify this claim, we plot the histograms of the Robinson $(1995a)$ estimator for $d = 1.4$, $m = n^{0.8}$ using four different orders $(p = 1, 2, 3, 4)$. The figure 1 shows that, the distribution seems to be normal only for $p = 2$, 3 and 4, i.e. for $p \geq [d + 1/2] + 1$.

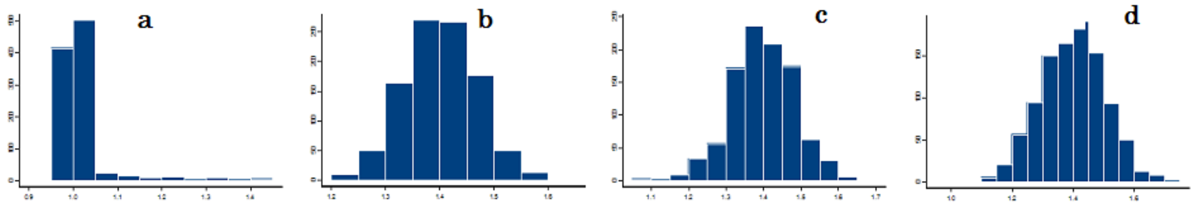

Figure 1: The histograms of the Robinson (1995a) estimator for $d = 1.4$, $m = n^{0.8}$. $a : p = 1$, $b : p = 2$, $c : p = 3$, $d : p = 4$.

- Given the previous comments, we conclude that the order p should be equal to $[d + 1/2] + 1$, and not higher, since the higher order yields large standard errors and also large MSE.

- Despite an optimal choice of p, the bias and the MSE are worse with $m = n^{0.5}$ than with $m = n^{0.8}$. Then, the truncation parameter m has an effect in the stationary as well as the nonstationary case.

4.2 An optimal choice of the truncation parameter m

In the stationary case, Hurvich et al. (1998) argued that the optimal choice of m is $cn^{0.8}$. However our paper deals with nonstationary process, so m optimal will tend to be larger than as given in Hurvich et al. (1998). To determine an optimal choice for m, we apply the heuristic method proposed by Abadir et al. (2005) and can be described as follows:

We assume that m has the form $m = cn^{\delta}$. The optimal choice for m can be achieved by an optimal choice for (c, δ). We can find c and δ by searching for the min MSE as function of the long memory parameter:

1. In the first step, we choose a grid for the sample size n, $G_1 = (100, 200, 400, 600, 800)$ and a grid for d, $G_2 = (0.6, 0.8, 1.2, 1.6)$. For n fixed in G_1 and for each d in G_2, we choose some values for the parameter m, $\left(m_n^1, m_n^2, ..., m_n^k\right)$. For each m_n^i, $1 \leq i \leq k$,[15] we compute the MSE of the corresponding $\hat{d}$: $MSE\left(m_n^i\right)$. We then compute (see table 7):[16]
$$\hat{m}_n = \min_{1 \leq i \leq k, d \in G_2} MSE\left(m_n^i\right),$$

2. In the second step, we regress $\log(\hat{m}_n)$ on $\log(n)$ in the following regression

[15] The choice of k is arbitrary.

[16] To save space, only the results for sample sizes $n = 100$ and $n = 800$ are given here, the others are supplied upon request.

$$\log(\hat{m}_n) = \log c + \delta \log(n) + errors.$$

The result indicates that $\hat{c} = 1$ and $\hat{\delta} = 0.868$, so the optimal choice of m is given by $\hat{m} = n^{0.868}$.

Table 7 The heuristic method for the choice of m

	The parameter m, n=100						The parameter m, n=800					
d	20	25	30	35	40	45	60	150	250	300	350	390
0.6	0.0836	0.0508	0.0405	0.0322	0.0274	0.0237	0.0143	0.0049	0.0029	0.0025	0.0022	0.0021
0.8	0.0833	0.0506	0.0404	0.0322	0.0273	0.0237	0.0142	0.0049	0.0029	0.0025	0.0022	0.0021
1.2	0.0851	0.0520	0.0415	0.0333	0.0283	0.0248	0.0243	0.0072	0.0043	0.0037	0.0032	0.0030
1.6	0.1729	0.1006	0.0911	0.0681	0.0538	0.0492	0.0249	0.0074	0.0044	0.0038	0.0033	0.0031

It may be seen from the results above that, in order to obtain a semiparametric estimator with good properties, one may choose the optimal pair (p, m). To see this, we plot, for $m = n^{0.5}$, $n^{0.55}$, $n^{0.6}$, $n^{0.65}$, $n^{0.7}$, $n^{0.75}$, $n^{0.8}$, $n^{0.868}$,[17] $p = 1, 2, 3, 4$ and $d = 1.2$, the MSE of the three semiparametric estimators.

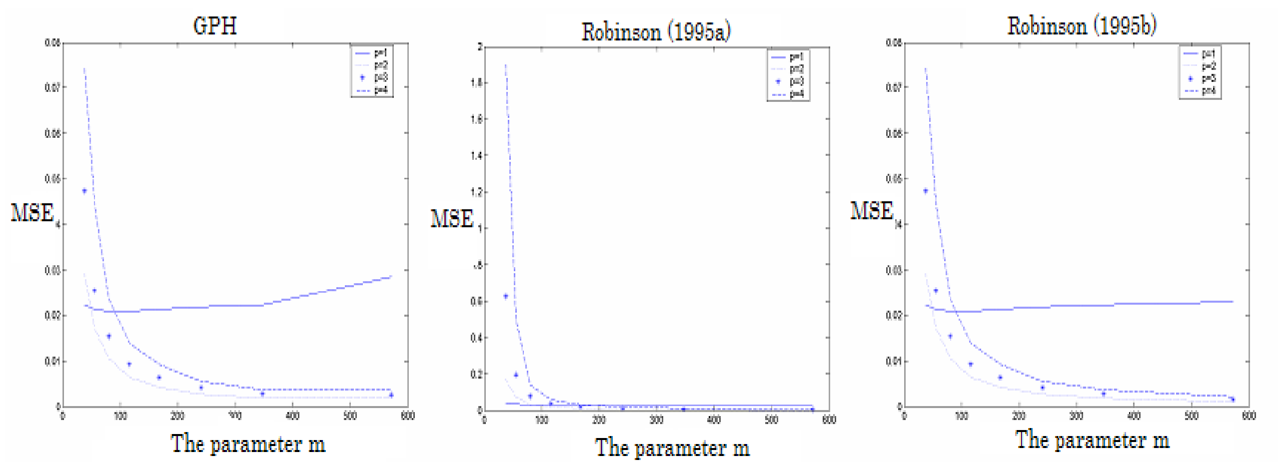

Figure 2: The MSE of the three semiparametric estimators for different values of m and for $ARFIMA\,(0, d, 0)$ process.

This figure shows that, if $p = 1$ $(p < [d + 1/2] + 1)$, the MSE vary slowly with m. Whereas if $p \geq [d + 1/2] + 1$, the smallest MSE is obtained for $m = n^{0.868}$.[18] The comparison

[17] $m = n^{0.5}$ is the choice of GPH (1983), $n^{0.55}$, $n^{0.6}$, $n^{0.65}$, $n^{0.7}$ and $n^{0.75}$ were suggested by Porter-Hudak (1990) and Crato and De Lima (1994), $m = n^{0.8}$ is the choice proposed by Hurvich et al. (1998) and finally $m = n^{0.868}$ is determined by the heuristic method suggested by Abadir et al. (2005).

[18] This choice of m is most preferable in our design.

of the MSE for $p = 2$, 3 and 4, shows that the smallest one is obtained with $p = 2$ which is equal to $[d + 1/2] + 1$.

4.3 The analysis of the ARFIMA (1,d,0) and ARFIMA (0,d,1) results

Now, we consider the presence of the short-memory component. Only the results of the Robinson (1995b) estimator are reported here. The same conclusions are obtained for the GPH and Robinson (1995a) estimators. They are supplied upon request.

- For almost all cases, the estimator is negatively biased, with bias approximately equal to $p-d$, when $p < [d + 1/2]+1$. Whereas it is positively biased when $p \geq [d + 1/2]+1$.
- The short-memory components ϕ and θ have an important impact on the results, even if we use the best choice of p. When ϕ and θ approach 1 in absolute value, the bias and the MSE become worse.
- With a good choice of p, the estimates for $ARFIMA$ $(0, d, 1)$ processes have nice properties than for $ARFIMA$ $(1, d, 0)$ processes. This leads us to conclude that, the short-memory component ϕ has a larger negative effect on the properties of semiparametric estimators than the parameter θ of same size.

Table 8 The Robinson $(1995b)$ estimator with "Zhurbenko Kolmogorov" taper: $ARFIMA\ (1, d, 0)$ with $\sigma = 1$.

$m = n^{0.5}$ and $n = 1500$													
		$p = 1$			$p = 2$			$p = 3$			$p = 4$		
d	ϕ	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE
0.6	0.40	0.616	0.097	0.010	0.599	0.164	0.027	0.597	0.216	0.047	0.590	0.276	0.076
	0.80	0.670	0.097	0.014	0.662	0.164	0.031	0.667	0.217	0.052	0.668	0.277	0.081
	0.90	0.794	0.097	0.047	0.808	0.166	0.071	0.828	0.220	0.100	0.843	0.280	0.137
	0.95	0.994	0.104	0.166	1.054	0.170	0.235	1.089	0.224	0.289	1.117	0.282	0.348
1.2	0.40	1.084	0.093	0.022	1.227	0.169	0.029	1.220	0.217	0.047	1.212	0.273	0.075
	0.80	1.102	0.104	0.020	1.289	0.169	0.036	1.290	0.217	0.055	1.290	0.274	0.083
	0.90	1.136	0.134	0.022	1.434	0.169	0.083	1.450	0.219	0.111	1.465	0.277	0.147
	0.95	1.176	0.183	0.034	1.678	0.172	0.258	1.711	0.223	0.311	1.740	0.280	0.369
1.8	0.40	1.086	0.151	0.532	1.887	0.165	0.035	1.860	0.219	0.052	1.848	0.271	0.076
	0.80	1.087	0.155	0.532	1.942	0.163	0.046	1.929	0.219	0.065	1.925	0.272	0.090
	0.90	1.090	0.164	0.531	2.060	0.159	0.093	2.088	0.220	0.131	2.098	0.275	0.165
	0.95	1.091	0.172	0.534	2.229	0.166	0.212	2.347	0.223	0.349	2.374	0.277	0.406

Table 9 The Robinson $(1995b)$ estimator with "Zhurbenko Kolmogorov" taper: $ARFIMA\ (1, d, 0)$ with $\sigma = 1$.

$m = n^{0.8}$ and $n = 1500$													
		$p = 1$			$p = 2$			$p = 3$			$p = 4$		
d	ϕ	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE
0.6	0.40	0.765	0.031	0.028	0.762	0.044	0.030	0.767	0.056	0.031	0.768	0.063	0.032
	0.80	1.188	0.044	0.348	1.218	0.051	0.385	1.224	0.063	0.394	1.230	0.072	0.402
	0.90	1.313	0.085	0.515	1.399	0.050	0.642	1.405	0.062	0.652	1.410	0.070	0.661
	0.95	1.321	0.131	0.537	1.498	0.047	0.809	1.503	0.059	0.818	1.507	0.066	0.828
1.2	0.40	1.095	0.102	0.024	1.366	0.044	0.026	1.368	0.055	0.031	1.370	0.063	0.033
	0.80	1.130	0.181	0.037	1.820	0.051	0.386	1.825	0.063	0.395	1.831	0.071	0.403
	0.90	1.120	0.182	0.040	2.000	0.050	0.643	2.006	0.062	0.653	2.011	0.070	0.663
	0.95	1.099	0.160	0.036	2.097	0.047	0.807	2.104	0.058	0.821	2.110	0.066	0.832
1.8	0.40	1.032	0.088	0.598	1.974	0.043	0.032	1.972	0.055	0.033	1.975	0.062	0.035
	0.80	1.032	0.093	0.598	2.293	0.112	0.256	2.427	0.063	0.397	2.433	0.071	0.406
	0.90	1.030	0.081	0.598	2.333	0.728	0.314	2.608	0.061	0.657	2.614	0.069	0.667
	0.95	1.031	0.096	0.599	2.294	0.200	0.284	2.707	0.058	0.827	2.713	0.066	0.838

Table 10 The Robinson $(1995b)$ estimator with "Zhurbenko Kolmogorov" taper: $ARFIMA\ (0, d, 1)$ with $\sigma = 1$.

$m = n^{0.5}$ and $n = 1500$													
		$p = 1$			$p = 2$			$p = 3$			$p = 4$		
d	θ	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE
0.6	0.40	0.608	0.097	0.009	0.590	0.163	0.027	0.587	0.216	0.047	0.579	0.275	0.076
	0.80	0.554	0.098	0.011	0.528	0.164	0.032	0.517	0.216	0.053	0.502	0.275	0.085
	0.90	0.430	0.100	0.038	0.387	0.165	0.072	0.360	0.217	0.104	0.330	0.275	0.148
	0.95	0.233	0.104	0.145	0.161	0.169	0.221	1.113	0.219	0.285	0.065	0.277	0.362
1.2	0.40	1.057	0.130	0.037	1.219	0.169	0.029	1.211	0.217	0.046	1.202	0.273	0.074
	0.80	1.059	0.123	0.036	1.157	0.169	0.030	1.141	0.217	0.050	1.124	0.272	0.080
	0.90	1.057	0.126	0.036	1.017	0.172	0.063	0.985	0.218	0.093	0.953	0.272	0.135
	0.95	1.061	0.135	0.037	0.792	0.175	0.197	0.738	0.220	0.262	0.687	0.274	0.338
1.8	0.40	1.081	0.144	0.537	1.878	0.166	0.034	1.850	0.219	0.050	1.837	0.271	0.075
	0.80	1.070	0.133	0.550	1.824	0.168	0.029	1.781	0.219	0.048	1.760	0.271	0.075
	0.90	1.044	0.100	0.580	1.695	0.174	0.041	1.627	0.221	0.079	1.589	0.271	0.118
	0.95	1.020	0.068	0.612	1.477	0.180	0.136	1.380	0.225	0.227	1.323	0.274	0.302

Table 11 The Robinson $(1995b)$ estimator with "Zhurbenko Kolmogorov" taper: $ARFIMA\ (0, d, 1)$ with $\sigma = 1$.

$m = n^{0.8}$ and $n = 1500$													
		$p = 1$			$p = 2$			$p = 3$			$p = 4$		
d	θ	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE	$\hat{d}^p_{Rb}$	$S\hat{E}$	MSE
0.6	0.40	0.466	0.033	0.019	0.449	0.044	0.025	0.443	0.054	0.027	0.437	0.063	0.030
	0.80	0.149	0.044	0.205	0.108	0.055	0.245	0.084	0.064	0.269	0.065	0.074	0.291
	0.90	−0.011	0.049	0.375	−0.064	0.060	0.445	−0.096	0.067	0.489	−0.121	0.076	0.526
	0.95	−0.146	0.051	0.559	−0.205	0.059	0.652	−0.238	0.064	0.707	−0.262	0.071	0.749
1.2	0.40	1.017	0.055	0.036	1.057	0.046	0.022	1.048	0.055	0.026	1.042	0.063	0.028
	0.80	1.016	0.056	0.036	0.721	0.059	0.233	0.694	0.066	0.260	0.673	0.074	0.283
	0.90	1.017	0.054	0.036	0.549	0.064	0.427	0.514	0.069	0.476	0.486	0.076	0.515
	0.95	1.020	0.060	0.031	0.408	0.063	0.631	0.370	0.066	0.692	0.344	0.072	0.738
1.8	0.40	1.029	0.078	0.600	1.690	0.057	0.015	1.659	0.056	0.023	1.651	0.063	0.026
	0.80	1.028	0.053	0.608	1.362	0.070	0.196	1.310	0.069	0.245	1.285	0.075	0.270
	0.90	1.011	0.029	0.623	1.191	0.074	0.376	1.131	0.073	0.453	1.099	0.078	0.497
	0.95	1.003	0.013	0.635	1.046	0.072	0.573	0.986	0.070	0.667	0.955	0.073	0.719

- The results show that with $m = n^{0.8}$ the bias is larger than with $m = n^{0.5}$ and the MSE is smaller only for $p \geq [d + 1/2] + 1$. Therefore we conclude that in the presence of short-memory parameter, m must be smaller than $n^{0.8}$. To verify this claim, we apply again the heuristic method proposed by Abadir et al. (2005) and described above. The results show that with an AR component, an optimal m is equal to $n^{0.597}$, whereas with MA component, it is equal to $n^{0.605}$.[19] In order to justify that, we plot for $m = n^{0.5}$, $n^{0.55}$, $n^{0.6}$, $n^{optimal}$, $n^{0.65}$, $n^{0.7}$, $n^{0.75}$, $n^{0.8}$, $p = 1, 2, 3, 4$ and $d = 1.2$, the MSE of the Robinson $(1995b)$ estimator, and this for $ARFIMA\ (1, d, 0)$ and $ARFIMA\ (0, d, 1)$ processes.

[19] The different results are supplied upon request.

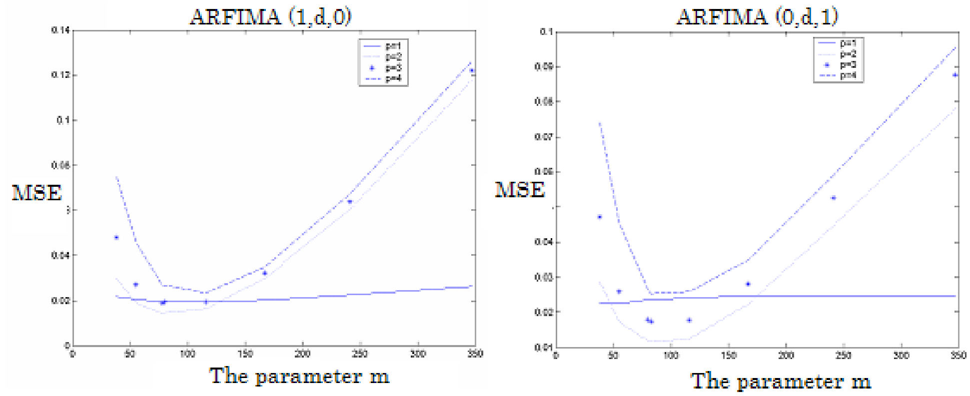

Figure 3: The MSE of the Robinson (1995b) estimator for different values of m and for $ARFIMA$ $(1, d, 0)$ and $ARFIMA$ $(0, d, 1)$ processes.

This figure exhibits that for $p \geq [d + 1/2] + 1$, the smallest MSE is obtained with $m = n^{optimal}$. We note here that our conclusion coincides, in the stationary case, with that of Hauser (1997).

5 Application to the CPI data

We consider a series of logarithm of Consumer Price Index (CPI)[20] in monthly frequencies, for four countries: France, Italy, Germany and U.S., over the period 1957 : 1 to 2002 : 3 ($n = 543$ observations). The data source is the IMF's International Financial Statistics. The following table gives some descriptive statistics.

Table 12 The descriptive Statistics

	France	Italy	Germany	U.S.
Mean	1.602	1.407	1.775	1.662
Maximum	2.040	2.071	2.046	2.069
Minimum	0.995	0.737	1.470	1.260
Variance	0.120	0.242	0.031	0.082
E-Kurtosis	1.408	1.330	1.551	1.406
Skewness	-0.177	-0.024	-0.187	-0.054
J.B	60.144	63.124	50.658	57.776
p-value	(0.000)	(0.000)	(0.000)	(0.000)

J.B is the Jarque-Bera Normality test.

[20] The CPI is an aggregate index of different prices in different sectors of economics. The series are seasonally adjusted by x_{12} $ARIMA$ process.

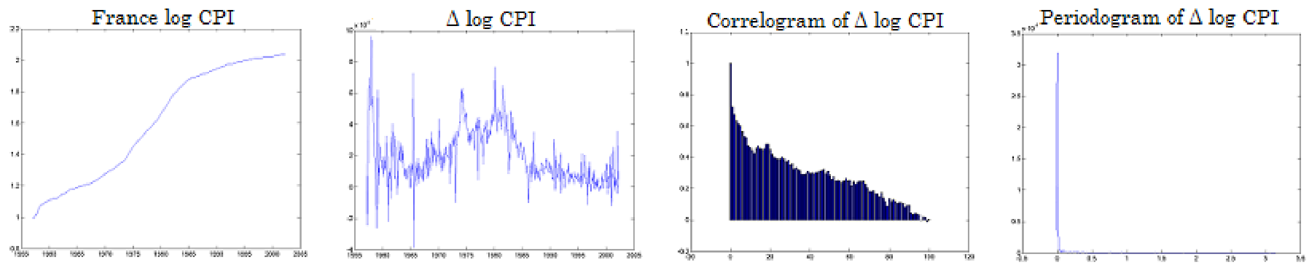

Figure 4: Graphs, Correlogram and Periodogram of the Original and First Differenced France logarithm CPI Series.

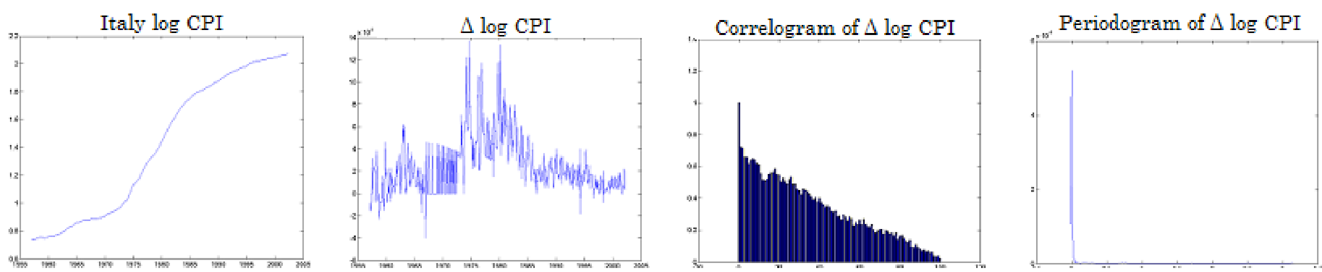

Figure 5: Graphs, Correlogram and Periodogram of the Original and First Differenced Italy logarithm CPI series.

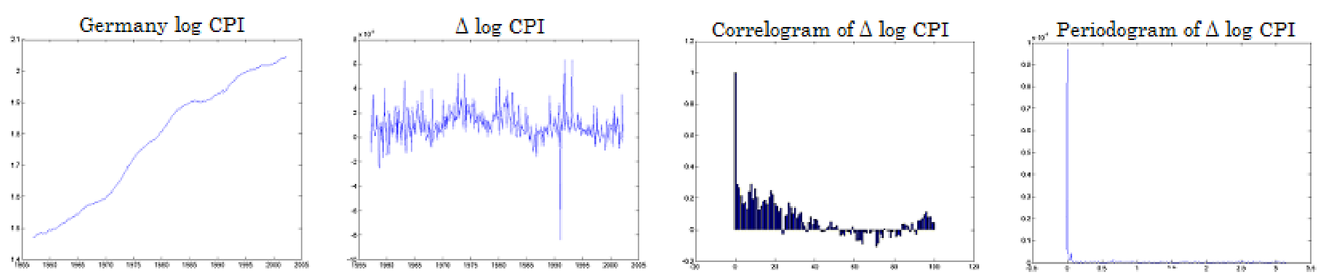

Figure 6: Graphs, Correlogram and Periodogram of the Original and First Differenced Germany logarithm CPI series.

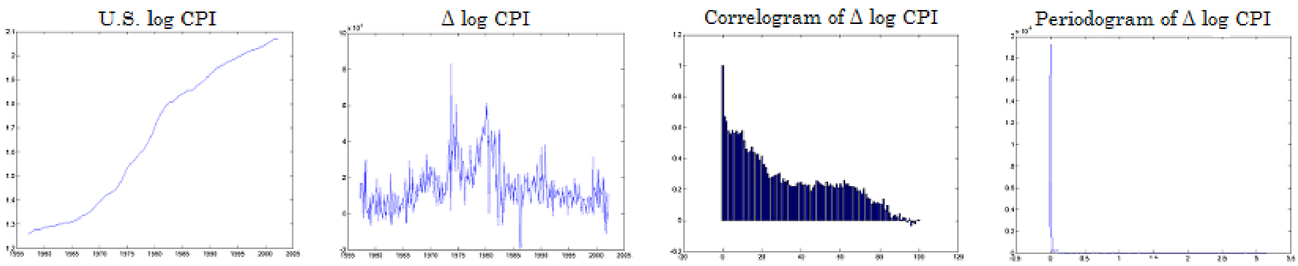

Figure 7: Graphs, Correlogram and Periodogram of the Original and First Differenced U.S.logarithm CPI series.

There is an overwhelming evidence that the fourth CPI series have non-normal distributions, as shown by the Jarque-Bera statistics. The correlograms and the periodograms of the series in first difference, (see figures 4, 5, 6 and 7), show a typical features of long-memory processes. The results of the stationarity tests[21] are given in table 13. Those tests are necessary to determine the order of differencing s and then the order of taper p. Usually, the estimation methods of d, especially the semiparametric ones are used only for $d < \frac{1}{2}$, however, as mentioned above, by using a taper with an adequate order p, we obtain a consistent and asymptotically normal estimator in the nonstationary case. By simulation, we showed that, the best choice of the order of taper is $\left[d + \frac{1}{2}\right] + 1$.

Table 13 Results of the stationarity tests

		ADF			P.P			KPSS		
		k			k			k		
		2	3	4	2	3	4	2	3	4
France	X_t	−2.241	−2.552	−2.241	2.716	2.328	2.045	2.697	2.026	1.624
	$\triangle X_t$	−6.311*	−5.086*	−4.240*	*−9.158	−9.295*	−9.513*	1.801	1.424	1.184
Italy	X_t	−0.309	−0.493	−0.607	−0.412	−0.469	−0.520	2.173	1.633	1.309
	$\triangle X_t$	−5.651*	−4.720*	−3.959*	−8.343*	−8.431*	−8.728*	2.690	2.097	1.723
Germany	X_t	−1.530	−1.452	−1.251	0.677	0.540	0.438	2.833	2.129	1.706
	$\triangle X_t$	−11.320*	−9.038*	−8.024*	−17.370*	−17.630*	−17.870*	0.834	0.723	0.648
U.S.	X_t	−0.038	−0.174	−0.261	−0.614	−0.656	−0.696	2.150	1.616	1.295
	$\triangle X_t$	−6.503*	−5.437*	−4.672*	−9.562*	−9.742*	−9.928*	2.164	1.709	1.422

* indicates that the hypothesis of stationarity is accepted.

Notes: 1. For the ADF test, k represents the number of lagged differences, and for the PP and KPSS tests, it is used for the construction of the spectral estimator.

2. For the ADF and PP tests, the critical values are -3.98(1%), -3.42(5%) and -3.13(10%), and for the KPSS one they are 0.216(1%), 0.146(5%) and 0.119(10%).

For different countries, the stationarity tests for the first differenced series give contradictory results, they show evidence of stationarity (PP and ADF) and unit root (KPSS).

[21] To test the stationarity, we use the Augmented Dickey-Fuller (ADF) (Dickey and Fuller (1979)), the Phillips and Perron (1988) (PP) and the KPSS (Kwiatkowski et al. (1992)) tests.

These contradictory statistical results coupled with the insights produced by examining the correlograms and the periodograms suggest that, a fractionally integrated model allowing for long memory is plausible for these once differenced series. These tests lead us to conclude that for the four series, $p = 2$, and then we apply the Robinson (1995*b*) method to determine the estimates of d.[22] The results are summarized in the following table.

Table 14 Results of the Robinson (1995*b*) estimators, $p = 2$.

m	France	Italy	Germany	U.S
$n^{0.5}$	2.142	1.807	1.742	2.226
$n^{0.55}$	1.951	1.770	1.713	2.075
$n^{0.60}$	1.893	1.887	1.556	1.881
$n^{0.65}$	1.793	1.864	1.446	1.671
$n^{0.70}$	1.652	1.606	1.339	1.581
$n^{0.75}$	1.646	1.564	1.368	1.569
$n^{0.80}$	1.618	1.588	1.339	1.603

These results show that $\hat{d}$ changes a lot with the parameter m, so, we can conclude that there is a serious short-run effects and hence the appropriate m should be chosen to be small. Based on figure 3, we see that $m^{optimal}$ is $n^{0.597}$ for an $ARFIMA\ (1, d, 0)$ process and $n^{0.605}$ for an $ARFIMA\ (0, d, 1)$ process. In our empirical application, we choose $m^{optimal} = n^{0.6}$, hence[23] $\hat{d} = 1.893$ for France, $\hat{d} = 1.887$ for Italy, $\hat{d} = 1.556$ for Germany and $\hat{d} = 1.881$ for U.S..

We can also verify the existence of short-run effects by testing the adequacy of the model. So, we apply a diagnostic tests to the residuals. First, we truncate the filter of equation (2.2) in the following way

$$(1 - B)^{\hat{d}} = \sum_{j=0}^{k} \hat{\tau}_j \left(\hat{d}\right) B^j, \qquad \hat{\tau}_j \left(\hat{d}\right) = \frac{j - 1 - \hat{d}}{j} \hat{\tau}_{j-1} \left(\hat{d}\right), \qquad \hat{\tau}_0 \left(\hat{d}\right) = 1. \tag{5.18}$$

According to Hassler and Wolters (1995), we choose k so that the following condition holds

$$\left|\hat{\tau}_{k-1} \left(\hat{d}\right)\right| \geq 0.005, \qquad \left|\hat{\tau}_k \left(\hat{d}\right)\right| < 0.005,$$

then, the residuals $\hat{\varepsilon}_t$ are given by

$$\hat{\varepsilon}_t = \sum_{j=0}^{k} \hat{\tau}_j \left(\hat{d}\right) X_{t-j}, \quad t = k + 1, ..., n. \tag{5.19}$$

[22] We will use different values of the parameter m ($n^{0.5}$, $n^{0.55}$, $n^{0.6}$, $n^{0.65}$, $n^{0.7}$, $n^{0.75}$, $n^{0.8}$).
[23] We choose $m^{optimal} = n^{0.6}$ because it is the value between $n^{0.597}$ and $n^{0.605}$.

Second, we check the residual autocorrelation by using the Ljung and Box (LB) (1978) statistic, and the normality by using the Shapiro-Wilk (SW) (1965) test. Table 15 indicated that the residuals are autocorrelated and non-normal, so, the model is not adequate.

To capture the autocorrelation, we fit an $ARMA(P,Q)$ process for $\hat{\varepsilon}_t$ by using the AIC criterion. The selected models for X_t for the four countries are given by

$$(1-B)^{1.893}(1-0.148B)\,CPI-FR_t=(1+0.713B)\,\varepsilon_t, \tag{5.20}$$

$$(1-B)^{1.887}(1+0.643B)\,CPI-IT_t=\left(1-0.054B+0.366B^2+0.102B^3\right)\varepsilon_t, \tag{5.21}$$

$$(1-B)^{1.556}(1-0.191B)\,CPI-GE_t=(1+0.534B)\,\varepsilon_t, \tag{5.22}$$

$$\begin{aligned}(1-B)^{1.881}(1+0.173B)\,CPI-US_t &= \Big(1+0.398B+0.121B^2+0.105B^3\\ &+ 0.061B^4-0.090B^5\Big)\,\varepsilon_t.\end{aligned} \tag{5.23}$$

The Ljung and Box (1978) statistics given in table 15,[24] confirm the absence of autocorrelation in the residuals of models ((5.20), (5.21), (5.22) and (5.23)). Furthermore, the CPI series for the four countries are generated by a nonstationary $ARFIMA\ (P,d,Q)$.

Table 15 Results of specification tests of original and transformed residuals, $m=n^{0.8}$.

Original residuals					Transformed residuals				
Test	France	Italy	Germany	U.S.	Test	France	Italy	Germany	U.S.
	86.787	111.787	51.602	93.021		6.636	0.191	0.267	0.010
LB	(0.000)	(0.000)	(0.000)	(0.000)	LB	(0.036)	(0.909)	(0.875)	(0.995)
	0.837	0.933	0.901	0.904		0.899	0.942	0.892	0.935
SW	(0.000)	(0.000)	(0.000)	(0.000)	SW	(0.000)	(0.000)	(0.000)	(0.000)

LB is the Ljung and Box test, SW is the Shapiro-Wilk test and the p-values are in the brakets.

6 Concluding Remarks

In this paper, we have studied by Monte Carlo simulations, the effect of order of "Zhurbenko-Kolmogorov" taper on the properties of three semiparametric estimators: the GPH, the Robinson (1995a) and the Robinson (1995b) estimators. According to the results we have observed that the order p of taper has an impact on the bias, standard errors and the mean squared errors of these three estimators. The smallest MSE is given by $p=[d+1/2]+1$

[24] See the results of transformed residuals.

which is considered as an optimal choice. When $p < [d + 1/2] + 1$, the d estimates are always negatively biased and converge to p.

Besides, the optimal choice of p, one may also choose a best value of m, since for a pure fractionaly integrated processes, the results obtained with $m = n^{0.5}$ have large bias and MSE and they improved with $m = n^{0.8}$. Whereas, for models with jointly short and long-run effects, the choice of m must be small.

We conclude that, to get a semiparametric estimators with small bias and MSE, we must choose an optimal pair (p, m). The short-memory component has also an effect on the properties of these estimators, so when ϕ and θ approach 1, the bias and MSE become larger, even if we have an appropriate choice of p and m. The comparison of the results of three estimators showed that the best ones are obtained for Robinson (1995b) estimator.

Acknowledgements: The authors would like to thank the editor and the anonymous referee for helpful comments and valuable suggestions. They also thank Philipp Sibbertsen for his answers on some questions about the procedure of tapering and Karim M. Abadir for the discussion of this paper.

References

Abadir A, Distaso W, Giraitis L (2005) Two estimators of the long-run variance. Discussion paper series 05/09, the university of York.

Agiakloglou C, Newbold P, Wohar M (1993) Bias in an estimator of the fractional difference parameter. Journal of Time Series Analysis **14**, 235 − 246.

Alekseev VG (1996) Jackson and Jackson-Vallée Poussin-type Kernels and Their Probability Applications. Theory of Probability and its Applications **41**, 137 − 143.

Beran J (1994) Statistics For Long-Memory Processes. Chapman & Hall.

Boutahar M, Marimoutou V, Nouira L (2006) Estimation methods of the long memory parameter : Monte Carlo analysis and application. Forthcoming in Journal of Applied Statistics.

Brillinger DR (1975) Time Series, Data Analysis and Theory. San Francisco, Holden-Day.

Cheung YW, Lai KS (1993) A Fractional Cointegration Analysis of Purchaising Power Parity. Journal of Business and Economic Statistics **11**, 103 − 112.

Cooley JW, Tukey JW (1965) An algorithm for the machine calculation of complex Fourier series. Mathematics of computation **19**, 297 − 301.

Cooley JW, Lewis PAW, Welch PD (1967) Historical notes on the fast Fourier transform. I.E.E.E. Trans. On Audio and Electroacoustics **15**, 76 − 79.

Crato N, De Lima PJF (1994) Long-range dependence in the conditional variance of stocks returns. Economics Letters **45**, 281 − 285.

Dickey D, Fuller W (1979) Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association **74**, 427 − 431.

Geweke J, Porter-Hudak S (1983) The Estimation And Application Of Long-Memory Time Series Models. Journal of Time Series Analysis **4**, 221 − 238.

Giraitis L, Robinson PM (2003) Edgeworth Expansions for Semiparametric Whittle Estimation of Long Memory. Annals of Statistics **31**, 221 − 238.

Granger CWJ, Joyeux R (1980) An introduction to long-memory time series models and fractional differencing. Journal of Time Series Analysis **1**, 1325 − 1375.

Hassler U, Wolters J (1995) Long memory in inflation rates: International evidence. Journal of Business and Economic Statistics **13**, 37 − 45.

Hauser MA (1997) Semiparametric and Nonparametric Testing for Long Memory: A Monte Carlo Study. Empirical Economics **22**, 247 − 271.

Hosking JRM (1981) Fractional Differencing. Biometrika **68**, 165 − 176.

Hurvich CM, Ray B (1995) Estimation of the memory parameter for nonstationary or noninvertable fractionally integrated processes. Journal of Time Series Analysis **16**, 17 − 42.

Hurvich CM, Deo RS, Brodsky J (1998) The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. Journal of Time Series Analysis **1**, 19 − 46.

Hurvich CM, Chen WW (2000) An efficient Taper for potentially over differenced Long-Memory Time Series. Journal of Time Series Analysis **21**, 155 − 180.

Jones RH (1971) Spectrum estimation with missing observations. Ann. Inst. Statist., Math. **23**, 387 − 398.

Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y (1992) Testing the hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root?. Journal of Econometrics **54**, 159 − 178.

Kim CS, Phillips PCB (1999) Log periodogram regression: the nonstationary case. Mimeographed, Cowles Foundation, Yale University.

Lobato I, Robinson PM (1996) Averaged periodogram estimation of long memory. Journal of Econometrics **73**, 303 − 324.

Ljung GM, Box GEP (1978) On a measure of lack of fit in time series models. Biometrika **65**, 553 − 564.

Maynard A, Phillips PCB (2001) Rethinking an Old Empirical Puzzle: Econometric Evidence on the Forward Discount Anomaly. Journal of Applied Econometrics **16**, 671 − 708.

Olhede SC, Mccoy EJ, Stephens DA (2004) Large-sample properties of the periodogram estimator of seasonally persistent processes. Biometrika **91**, 613 − 628.

Phillips PCB, Perron P (1988) Testing for a unit root in time series regression. Biometrika **75**, 335 − 346.

Phillips PCB (2005) Econometric Analysis of Fisher's Equation. American Journal of Economics and sociology **64**, 125 − 168.

Porter-Hudak S (1990) An Application of the Seasonal Fractionally Differenced Model to the Monetary Aggregates. Journal of the American Statistical Association **85**, 338 − 344.

Reisen VA (1994) Estimation of the fractional difference parameter in the $ARIMA\,(p, d, q)$ model using the smoothed periodogram. Journal of Time Series Analysis **15**, 335 − 350.

Robinson PM (1994) Semiparametric analysis of long-memory time series. The Annals of Statistics **22**, 515 − 539.

Robinson PM (1995*a*) Log-Periodogram Regression of Time Series With Long Range Dependence. The Annals of Statistics **23**, 1048 − 1072.

Robinson PM (1995*b*) Gaussian Semiparametric Estimation of Long Range Dependence. The Annals of Statistics **23**, 1630 − 1661.

Shapiro SS, Wilk MB (1965) An Analysis of Variance Test For Normality. Biometrika **52**, 591 − 611.

Shimotsu K, Phillips PCB (2004) Exact Local Whittle Estimation of Fractional Integration. The Annals of Statistics **32**, 656 − 692.

Sibbertsen P (2004) Long-memory in volatilities of German stock returns. Empirical Economics **29**, 477 − 488.

Taqqu MS, Teverovsky V (1996) Semiparametric graphical estimation techniques for long memory data. Lecture Notes in Statistics **115**, 420 − 432.

Velasco C (1999*a*) Non-stationary log-periodogram regression. Journal of Econometrics **91**, 325 − 371.

Velasco C (1999*b*) Gaussian semiparametric estimation of non-stationary time series. Journal of Time Series Analysis **20**, 87 − 127.

Zhurbenko IG (1979) On the Efficiency of Estimates of a Spectral Density. Scandinavian Journal of Statistics **6**, 49 − 56.