Journal of Applied Statistics
Vol. 34, No. 3, 261–301, April 2007

Routledge
Taylor & Francis Group

Estimation Methods of the Long Memory Parameter: Monte Carlo Analysis and Application

MOHAMED BOUTAHAR, VÊLAYOUDOM MARIMOUTOU & LEÏLA NOUIRA

GREQAM, Université de la Méditerranée, Marseille, France

ABSTRACT *Since the seminal paper of Granger & Joyeux (1980), the concept of a long memory has focused the attention of many statisticians and econometricians trying to model and measure the persistence of stationary processes. Many methods for estimating d, the long-range dependence parameter, have been suggested since the work of Hurst (1951). They can be summarized in three classes: the heuristic methods, the semi-parametric methods and the maximum likelihood methods. In this paper, we try by simulation, to verify the two main properties of* $\hat{d}$*: the consistency and the asymptotic normality. Hence, it is very important for practitioners to compare the performance of the various classes of estimators. The results indicate that only the semi-parametric and the maximum likelihood methods can give good estimators. They also suggest that the AR component of the ARFIMA* $(1, d, 0)$ *process has an important impact on the properties of the different estimators and that the Whittle method is the best one, since it has the small mean squared error. We finally carry out an empirical application using the monthly seasonally adjusted US Inflation series, in order to illustrate the usefulness of the different estimation methods in the context of using real data.*

KEY WORDS: Long memory, ARFIMA (p, d, q) process, fractional Gaussian noise, Monte Carlo study

Introduction

The presence of strong dependence in time series was highlighted by Newcomb (1886) and Jeffreys (1939) using astronomical data and by Student (1927) using an industrial production data. Hurst (1951) established an empirical law, and built a test to detect a long memory in hydrological data. In economics, the long memory component was first detected in exchange rate data (Mandelbrot, 1962; Cheung, 1993; Beran & Ocker, 1999; Velasco, 1999), in stock prices data (Cheung & Lai, 1993; Chow *et al.*, 1995; Bhardwaj & Swanson, 2006) in macroeconomics data (Hassler & Wolters, 1995; Hyung & Franses, 2001; Bos *et al.*, 2002; Chio & Zivot, 2002; and Stock & Watson, 2002).

Correspondence Address: Leïla Nouira, GREQAM, Université de la Méditerranée, Centre de la Vieille Charité, 2 Rue de la Charité, 13002 Marseille, France. Email: leila.nouira@univmed.fr

0266-4763 Print/1360-0532 Online/07/030261–41
DOI: 10.1080/02664760601004874

A stationary process $\{X_t\}$ is called a long-range dependence or long memory process if

- there exist $\alpha \in (0, 1)$ and a constant $c_\rho > 0$ such that $\lim_{k\to\infty} k^\alpha \rho(k) = c_\rho$, where $\rho(k)$ is the autocorrelation function, or
- there exist $\beta \in (0, 1)$ and a constant $c_f > 0$ such that $\lim_{\lambda\to 0} |\lambda|^\beta f(\lambda) = c_f$, where $f(\lambda)$ is the spectral density function.

Among the models that can check these two definitions, we can consider the fractional Gaussian noise (further referred to as FGN), introduced at the beginning of the 1970s by Mandelbrot & Van Ness (1968) and Mandelbrot (1971), and the ARFIMA (p, d, q) process, introduced by Granger & Joyeux (1980) and Hosking (1981), which generalizes the Box & Jenkins (1976) ARIMA (p, d, q) process by allowing d to take real values. The link between the self-similarity parameter H of FGN and the ARFIMA parameter d is that $H = d + 1/2$, this relation is obtained by using the behaviour of their spectral densities (see equations (11) and (18)).[1]

The properties of the process $\{X_t\}$ depend on the value of the parameter d. Many researchers, such as, Lo, Mandelbrot, Künsch, Jensen, Taqqu, Whittle, Geweke & Porter-Hudak, Robinson, and Reisen among many others, have proposed methods for estimating the self-similarity parameter H or the long memory parameter d. These methods can be summarized in three classes: the heuristic methods (e.g. Hurst, 1951; Higuchi, 1988; Lo, 1991), the semiparametric methods (e.g. Geweke & Porter-Hudak, 1983; Robinson, 1994, 1995a and 1995b; Reisen, 1994; Lobato & Robinson, 1996) and the maximum likelihood methods (e.g. Whittle, 1951; Sowell, 1992). In the first two classes, we can estimate only the long memory parameter d. However, to fit an ARFIMA (p, d, q) model, we need two steps: one filters out the long memory component and then fits an ARMA (p, q) model to the residual series. In the last class, we estimate simultaneously all the parameters. It is important to note that the FGN and the ARFIMA process belong to two different classes of models, so the first is a self-similar process, whereas the second isn't, which makes the comparison of these two models uneasy. However, if they are Gaussian and if the ARFIMA (p, d, q) process is canonical, then the estimators of these two models are comparable (Beran, 1994; Taqqu & Teverovsky, 1998).

The asymptotic properties of the different estimators have been studied theoretically only for some of them. For example, for the heuristic methods, the researchers have suggested the estimators without studying the asymptotic properties. The aim of this paper is first, to show empirically the consistency and the asymptotic normality of some estimators and second to compare the performance of the various classes of estimators.

In the next section, we will present, briefly, the two families of processes. In the section after we will describe the different estimators $\hat{d}$ to be compared and give their asymptotic properties. In the fourth section, we will perform the Monte Carlo simulations. An empirical application on the behaviour of the monthly seasonally adjusted US Inflation rate, will be presented in the fifth section.

Definitions and Characteristics of ARFIMA (p, d, q) Process and Fractional Gaussian Noise

The ARFIMA (p, d, q) Process

In order to model a time series with long memory behaviour, Granger & Joyeux (1980) and Hosking (1981) proposed a class of models, the ARFIMA (p, d, q) process, where the parameter d can take real values. $\{X_t\}$ is a canonical ARFIMA (p, d, q) process, if it is a solution of

$$\phi(B)(1-B)^d(X_t-\mu)=\theta(B)\varepsilon_t, \tag{1}$$

with $\phi(B)=1-\phi_1 B-\cdots-\phi_p B^p, \theta(B)=1+\theta_1 B+\cdots+\theta_q B^q, d\in\Re$, B is a lag operator, $\mu=E(X_t)$, $\varepsilon_t\sim i.i.d.(0,\sigma_\varepsilon^2)$ and

$$(1-B)^d=\sum_{k\geq 0} b_k(d)B^k \tag{2}$$

where

$$b_k(d)=\frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)} \quad \text{for} \quad k\geq 0 \tag{3}$$

$\Gamma(\cdot)$ is the gamma function. The spectral density of $\{X_t\}$ is

$$f(\lambda)=\frac{\sigma_\varepsilon^2}{2\pi}\frac{|\theta(e^{i\lambda})|^2}{|\phi(e^{i\lambda})|^2}|1-e^{i\lambda}|^{-2d} \quad \text{for} \quad -\pi<\lambda<\pi \tag{4}$$

$f(\lambda)$ is continuously differentiable for all non-zero frequencies. For $d>0$, $f(\lambda)$ is discontinuous and unbounded at zero.

- if $d>-1/2$, then $\{X_t\}$ is invertible,
- if $d<1/2$, then $\{X_t\}$ is stationary,
- if $-1/2<d<0$, then $\rho(k)$ decreases more quickly than the case $0<d<1/2$. There is a stronger mean reversion, and $\{X_t\}$ is called anti-persistent in Mandelbrot's terminology,
- if $0<d<1/2$, $\{X_t\}$ is a stationary long memory process. The autocorrelation function decays hyperbolically to zero and we have

for $|k|\to\infty$, $\quad \gamma(k)\equiv C_\gamma(d,\phi,\theta)|k|^{2d-1}$ where

$$C_\gamma(d,\phi,\theta)=\frac{\sigma_\varepsilon^2}{\pi}\frac{|\theta(1)|^2}{|\phi(1)|^2}\Gamma(1-2d)\sin\pi d \tag{5}$$

and then

$$\text{for} \quad |k|\longrightarrow\infty, \quad \rho(k)=\frac{\gamma(k)}{\gamma(0)}\equiv\frac{C_\gamma(d,\phi,\theta)}{\int_{-\pi}^{\pi}f(\lambda)\,\mathrm{d}\lambda}|k|^{2d-1}$$

- if $d=1/2$, then at zero frequency the spectral density is unbounded.

Now, we consider the ARFIMA $(0,d,0)$ process

$$(1-B)^d(X_t-\mu)=\varepsilon_t, \quad \varepsilon_t\sim i.i.d.(0,\sigma_\varepsilon^2) \tag{6}$$

If $d>-1/2$, then the process is invertible and has an infinite autoregressive representation

$$\pi(B)X_t=\sum_{k=0}^{\infty}\pi_k X_{t-k}=\varepsilon_t \tag{7}$$

where

$$\pi_k=\frac{\Gamma(k-d)}{\Gamma(-d)\,\Gamma(k+1)} \tag{8}$$

If $d < 1/2$, then $\{X_t\}$ is stationary and its moving average representation is

$$X_t = \theta(B)\varepsilon_t = \sum_{k=0}^{\infty} \theta_k \varepsilon_{t-k} \tag{9}$$

where

$$\theta_k = \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)} \tag{10}$$

If $-1/2 < d < 1/2$, then the process is invertible and stationary with

$$\begin{aligned} f(\lambda) &= \frac{\sigma_\varepsilon^2}{2\pi}\left(2\sin\frac{\lambda}{2}\right)^{-2d} \quad \text{for } 0 < \lambda \leq \pi, \\ &\equiv C\lambda^{-2d}, \quad \text{if } \lambda \longrightarrow 0 \text{ where } C \text{ is a constant,} \end{aligned} \tag{11}$$

where C is a constant and

$$\rho(k) = \frac{\Gamma(1-d)\Gamma(k+d)}{\Gamma(d)\Gamma(k-d+1)} \equiv \frac{\Gamma(1-d)}{\Gamma(d)} k^{2d-1}, \quad \text{if} \quad k \longrightarrow \infty \tag{12}$$

If $d = -1/2$, then the ARFIMA $(0, d, 0)$ is stationary and non-invertible with

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{\pi}\sin(\lambda/2), \quad \lim_{\lambda \to 0} f(\lambda) = 0 \tag{13}$$

and

$$\rho(k) = -\frac{1}{4k^2 - 1} \tag{14}$$

If $d = 0$, then the ARFIMA $(0, d, 0)$ is a white noise.

The Fractional Gaussian Noise

The best way to introduce the fractional Gaussian noise is to do it from the fractional Brownian motion $\{B_H(t), t \geq 0\}$. The fractional Brownian motion is a zero mean Gaussian process with stationary increments, variance $E(B_H^2(t)) = t^{2H}$ and covariance function

$$E(B_H(s)B_H(t)) = \frac{1}{2}\left\{s^{2H} + t^{2H} - |s-t|^{2H}\right\} \tag{15}$$

It is self-similar in the sense that $\{B_H(at), t \geq 0\}$ has the same distribution in finite dimension as $\{a^H B_H(t), t \geq 0\}$ for all $a > 0$. The important index is H, a parameter on [0, 1], it is called a parameter of self-similarity or a Hurst coefficient.[2] The fractional Gaussian noise $\{X_t, t \geq 1\}$ is the increment of the fractional Brownian motion

$$X_t = B_H(t) - B_H(t-1), \quad t \geq 1 \tag{16}$$

centred stationary Gaussian with autocovariance function

$$\begin{aligned} \gamma(k) &= E(X_t X_{t+k}) = \frac{1}{2}\{\{k+1\}^{2H} - 2k^{2H} + |k-1|^{2H}\}, \quad k \geq 0 \\ \gamma(k) &\sim \mathrm{H}(2\mathrm{H}-1)k^{2H-2} \quad \text{for } k \longrightarrow \infty \text{ and } H \neq \frac{1}{2} \end{aligned} \tag{17}$$

If $H = 1/2$, then $\gamma(k) = 0$ for all $k \geq 1$, consequently $\{X_t\}$ is a white noise process. The $\{X_t\}$ is positively correlated when $1/2 < H < 1$ and it is called a long memory process.

In this framework, H measures the intensity of long-range dependence. The spectral density is

$$f(\lambda) = C_H \left(2 \sin \frac{\lambda}{2}\right)^2 \sum_{k=-\infty}^{+\infty} \frac{1}{|\lambda + 2\pi k|^{2H+1}} \sim \mathrm{C_H}|\lambda|^{1-2H}$$

$$\text{when } \lambda \to 0, \quad C_H \text{ is a constant} \tag{18}$$

The advantage of a fractional Gaussian noise on an ARFIMA (p, d, q) process is that the asymptotic relations are checked for finite sample sizes, because the fractional Gaussian noise is an increment of a self-similar fractional Brownian motion. The advantage of an ARFIMA (p, d, q) on a fractional Gaussian noise is that, it has simpler spectral density (4). However, the two spectral densities have the same behaviour in $|\lambda|^{-2d}$ when $\lambda \to 0$ (with $H = d + 1/2$).

Estimation Methods of Long Memory Parameter

The Heuristic Methods

The Hurst method: the statistic R/S. This method is based on the statistic $Q(n) = R(n)/S_n$, with

$$R(n) = \max_{1\le k\le n} \sum_{i=1}^{k} \left(X_i - \overline{X_n}\right) - \min_{1\le k\le n} \sum_{i=1}^{k} \left(X_i - \overline{X_n}\right), \tag{19}$$

$$S_n^2 = n^{-1} \sum_{i=1}^{n} \left(X_i - \overline{X_n}\right)^2, \tag{20}$$

and

$$\overline{X_n} = n^{-1} \sum_{i=1}^{n} X_i, \tag{21}$$

where n is the sample size. This method allows detecting the non-periodic cycles; then we can estimate H.

In practice, this method is done in several steps:

- First, we determine a sequence of integers $(k_i)_{i=1,\ldots,m}$ with length m, arbitrarily chosen, such as $1 < k_m < \cdots < k_1 < n$, for which we use a sequence defined by Davies & Harte (1987) such as $\forall\, i = 1, 2, 3, \ldots, 6$, $k_i = [n/i]$ and $\forall\, i = 7, 8, 9, \ldots, m$, $k_i = [k_{i-1}/1.15^i]$, where $[\cdot]$ is the integer part.
- For each k_i, we compute the statistic $Q(k_i)$.
- Then, we apply the least squares method on the regression model $\log Q(k_i) = a + b \log(k_i) + u_i$. The slope estimate is the Hurst coefficient $\hat{H}$.

The only advantage of this method is that it gives the possibility to obtain an estimator with good properties of robustness (Mandelbrot & Taqqu, 1979). On the other hand, it has several disadvantages, in particular, the exact distribution of the statistic R/S is difficult to determine and depends on the distribution of the process.

The Lo method. One of the disadvantages of the R/S statistic is its sensitivity to the short-memory effect. To overcome this problem, Lo (1991) proposed the 'modified R/S statistic'.

Its limiting distribution is invariant irrespective to the various forms of the processes with short-memory. The modified R/S statistic has the following form

$$\tilde{Q}_q(n) = \frac{R(n)}{S_q(n)} \tag{22}$$

where

$$S_q(n) = \left\{ S_n^2 + \frac{2}{n} \sum_{j=1}^{q} w_j(q) \left[\sum_{i=j+1}^{n} \left(X_i - \bar{X}_n\right)\left(X_{i-j} - \bar{X}_n\right) \right] \right\}^{1/2}$$

S_n^2 and $\bar{X}_n$ are the empirical variance and the empirical mean defined respectively by equations (20) and (21), $w_j(q) = 1 - j/(q+1)$, are the weights proposed by Newey & West (1987), with $j = 1, \ldots, q$. Phillips (1987) showed the convergence in probability of $S_q(n)$ under the two following conditions

1. $\sup_t E\left\lfloor |X_t|^{2\beta} \right\rfloor < \infty$, $\beta > 2$,
2. If $n \to \infty$, then $\mathrm{q} \to \infty$ such that $\mathrm{q} \sim o(\mathrm{n}^{1/4})$.

There is no optimal choice of the parameter q. Lo & MacKinlay (1989) and Andrews (1991) showed by a Monte Carlo study that, when q is relatively large compared to the sample size, then the estimator is skewed and thus q must be relatively small.

The Higuchi method. Higuchi (1988) proposed a method that allows estimating the fractal dimension D of a non-periodic and irregular time series. For a self-similar processes, the fractal dimension is $D = 1 - H$. The method for estimating H with the sample $\{X_1, X_2, \ldots, X_n\}$, is the following

- For a fixed k, we build k sub-samples

$$X_k^l = \left\{X_l, X_{l+k}, X_{l+2k}, \ldots, X_{l+[n-l/k]k}\right\}, \quad l = 1, 2, \ldots, k, \quad \text{where} \quad k \le \frac{n}{64} \tag{23}$$

- We determine the length of the curve X_k^l with the following formula

$$L_l(k) = \frac{n-1}{[n-l/k]k^2} \sum_{i=1}^{[n-l/k]} \left|X_{l+ik} - X_{l+(i-1)k}\right| \tag{24}$$

- We calculate the length of all the curves, $L(k)$ as

$$L(k) = \frac{1}{k} \sum_{l=1}^{k} L_l(k) \tag{25}$$

- For a long memory process, the following relation holds

$$E(L(k)) \sim \mathrm{c_H} k^{-D} \tag{26}$$

Taking the log in equation (26), we have $\log(E(L(k))) = \log(c_H) - D\log(k) + \upsilon_k$. By projecting $\log E(L(k))$ on $\log(k)$, we obtain the estimate $\hat{D}$ and then we determine $\hat{H} = 1 - \hat{D}$.

There are several other heuristic methods, such as the methods based on the variances (aggregated variance, absolute values of the aggregated series, differencing the variance) introduced by Taqqu *et al.* (1995). The efficiency of these methods was analysed in Kurpiel & Marimoutou (2000). Other approaches such as the variogram, the correlogram (Beran, 1994), the wavelets (Jensen, 1994) can also be considered. Today, no asymptotic properties of the heuristic estimators (convergence and asymptotic normality) are known.

The Semi-parametric Methods

Geweke & Porter-Hudak (further referred as GPH) developed a semi-parametric method at the beginning of the 1980s. This method is based on the behaviour of the spectral density of ARFIMA (p, d, q) process when the frequencies tend towards zero. As previously, this method allows us to estimate only the long-memory parameter d. The GPH estimator presents a bias related to the periodogram estimator. Hurvich & Beltrao (1993) and Robinson (1994, 1995a) proposed a modified version of the GPH estimator by using a smoothed periodogram or by discarding the first frequencies to reduce this bias. In this section, we will be interested in the presentation of the GPH method and those proposed by Robinson (1995a,b).

The Geweke and Porter-Hudak method. To illustrate this method, we can write the spectral density of $\{X_t\}$ as

$$f(\lambda) = \left\{4\sin^2\left(\frac{\lambda}{2}\right)\right\}^{-d} f_\varepsilon(\lambda) \tag{27}$$

where $f_\varepsilon(\lambda)$ is the spectral density of $\varepsilon_t = (1-B)^d X_t$, assumed to be a finite and continuous function on the interval $[-\pi, \pi]$. It follows that

$$\log\{f(\lambda)\} = \log\{f_\varepsilon(0)\} - d\log\left\{4\sin^2\left(\frac{\lambda}{2}\right)\right\} + \log\left\{\frac{f_\varepsilon(\lambda)}{f_\varepsilon(0)}\right\} \tag{28}$$

Let $I\left(\lambda_{j,n}\right)$ be the periodogram evaluated at the Fourier frequencies $\lambda_{j,n} = 2\pi j/n$, $(j = 1, 2, \ldots, m)$, where m,[3] is the number of frequencies which will be used in the regression;

$$\log\{I(\lambda_{j,n})\} = \log\{f_\varepsilon(0)\} - d\log\left\{4\sin^2\left(\frac{\lambda_{j,n}}{2}\right)\right\} + \log\left\{\frac{f_\varepsilon\left(\lambda_{j,n}\right)}{f_\varepsilon(0)}\right\} + \log\left\{\frac{I\left(\lambda_{j,n}\right)}{f\left(\lambda_{j,n}\right)}\right\} \tag{29}$$

where $\log\{f_\varepsilon(0)\}$ is a constant, $\log\{4\sin^2(\lambda_{j,n}/2)\}$ is an exogenous variable and $\log\{I(\lambda_{j,n})/f(\lambda_{j,n})\}$ is a disturbance term. The GPH estimator requires two major assumptions:

- Assumption 1: for sufficiently low frequencies, $\log\{f_\varepsilon(\lambda_{j,n})/f_\varepsilon(0)\}$ is negligible.
- Assumption 2: the random variables $\log\{I(\lambda_{j,n})/f(\lambda_{j,n})\}_{j=1,\ldots,m}$ are asymptotically *i.i.d.*

Under these two assumptions, we can write the linear regression

$$\log\left\{I\left(\lambda_{j,n}\right)\right\} = \alpha - d\log\left\{4\sin^2\left(\frac{\lambda_{j,n}}{2}\right)\right\} + e_j \tag{30}$$

where $e_j \sim$ i.i.d.$(-c, (\pi^2/6))$.[4] Let $Y_j = -\log\{4\sin^2(\lambda_{j,n}/2)\}$, then the GPH estimator is

$$\hat{d}_{GPH} = \frac{\sum_{j=1}^{m}\left(Y_j - \bar{Y}\right)\log\left\{I\left(\lambda_{j,n}\right)\right\}}{\sum_{j=1}^{m}\left(Y_j - \bar{Y}\right)^2}, \quad \bar{Y} = \frac{1}{m}\sum_{j=1}^{m} Y_j \tag{31}$$

Geweke & Porter-Hudak (1983) showed the convergence in probability, the asymptotic normality for $d < 0$ and $m = g(n)$ with $\lim_{n\to\infty} g(n) = \infty$ and $\lim_{n\to\infty} g(n)/n = 0$. Porter-Hudak (1990), Crato and De Lima (1994) showed that the parameter m must be selected so that $m = n^{\upsilon}$, with $\upsilon = 0.5, 0.6, 0.7$.

The Robinson estimation methods. To mitigate the bias of the GPH estimator, Robinson (1995a) proposed an asymptotic unbiased estimator for d,[5] which doesn't take the l first frequencies. This estimator is given by

$$\hat{d}_{Ra} = \frac{\sum_{j=l+1}^{m}\left(Y_j - \bar{Y}\right)\log\left\{I\left(\lambda_{j,n}\right)\right\}}{\sum_{j=l+1}^{m}\left(Y_j - \bar{Y}\right)^2}, \quad 0 \le l < m < n \tag{32}$$

with

$$\bar{Y} = \frac{1}{m-l}\sum_{j=l+1}^{m} Y_j \tag{33}$$

Under the assumption of Gaussianity and some conditions of regularity, the estimator $\hat{d}_{Ra}$ is consistent and asymptotically normal. There is no optimal choice for the parameters l and m which cause an important problem in the implementation of this method, so a bad choice of these two parameters can increase the bias of the estimator.

In Robinson (1995b), we assume that the spectral density of the process satisfies:

$$f(\lambda) \sim \mathrm{G}\lambda^{1-2\mathrm{H}} \quad \text{if } \lambda \longrightarrow 0^+$$

To determine $\hat{H}$, let

$$R(H) = \log \hat{G}(H) - (2H-1)\frac{1}{m}\sum_{j=1}^{m}\log\lambda_{j,n} \tag{34}$$

where

$$\hat{G}(H) = \frac{1}{m}\sum_{j=1}^{m}\lambda_j^{2H-1} I\left(\lambda_{j,n}\right) \tag{35}$$

with $m = n/2$ for n even, and $m = (n+1)/2$ for n odd. The estimator $\hat{H}$ is determined as

$$\hat{H} = \arg\min_{H\in\Theta} R(H) \tag{36}$$

where $\Theta = [\Delta_1, \Delta_2]$, Δ_1 and Δ_2 are chosen arbitrarily between 0 and 1, such that $0 < \Delta_1 < \Delta_2 < 1$. Robinson (1995b) established the convergence in probability and the asymptotic normality of $\hat{H}$ under strong regularity conditions.

Note that $\hat{H}$ is the maximum likelihood estimator in the following function

$$f(\lambda) = G|\lambda|^{1-2H}, \quad \lambda \in (-\pi, \pi) \tag{37}$$

with $G \in]0, \infty[$ and $H \in [0, 1]$. Moreover, this estimator has a rate of convergence of $m^{1/2}$ and it is much less efficient than the one obtained from the correctly specified parametric model.

The Maximum Likelihood Methods

In the semi-parametric methods, we cannot estimate simultaneously the interest parameter $\theta \equiv (\sigma_\varepsilon^2, d, \phi_1, \phi_2, \ldots, \phi_p, \theta_1, \theta_2, \ldots, \theta_q)$, then we consider the most efficient methods to estimate it. Two approaches are considered: the exact and the approximate Whittle (1951) maximum likelihood methods.[6]

The exact maximum likelihood method. Let $\{X_n\}$ be a canonical ARFIMA (p, d, q) process with p and q fixed, $E(X_n) = \mu = 0$ and $d < 1/2$. We consider a sample $X = [X_1, X_2, \ldots, X_n]^T \in \Re^n$, with $X \sim \mathrm{N}\big(0, \sum_\mathrm{n}\big)$, $\sum_n$ is a Toeplitz matrix

$$\sum\nolimits_n = \big[\gamma(j-l)\big]_{j,l=1,2,\ldots,n} = [\mathrm{cov}(X_j, X_l)]_{j,l=1,2,\ldots,n} \tag{38}$$

The log likelihood function is given by

$$L_n(X,\theta) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log\left|\sum\nolimits_n\right| - \frac{1}{2}X^T\sum_n^{-1}X \tag{39}$$

where $\left|\sum_n\right|$ is the determinant of $\sum_n$. The maximum likelihood estimator $\hat{\theta}$ is

$$\hat{\theta} = \arg\max_{\theta\in\Theta} L_n(X,\theta) \tag{40}$$

Yajima (1985) showed for an ARFIMA $(0, d, 0)$, the convergence in probability and the asymptotic normality of $\hat{\theta}$ under strong regularity conditions. Thereafter, Dahlhaus (1989) established these two properties in the case of an ARFIMA (p, d, q), and also the efficiency of this estimator.

Since, in equation (39), we need to compute log $\left|\sum_n\right|$ and $\sum_n^{-1}$, this method generates a high computational cost. Then to overcome this problem, we consider the approximate maximum likelihood method suggested by Whittle (1951).

The Whittle approximate maximum likelihood method. The Whittle (1951) method proceeds in several steps:

- First, we approximate log $\left|\sum_n\right|$. Grenander & Szegö (1958) showed that

$$\lim_{n\to\infty}\frac{1}{n}\log\left|\sum\nolimits_n\right| = \frac{1}{2\pi}\int_{-\pi}^{\pi}\log f(\lambda,\theta)\,\mathrm{d}\lambda \tag{41}$$

  then we substitute in equation (39) log $\left|\sum_n\right|$ by $n/2\pi \int_{-\pi}^{\pi} \log f(\lambda,\theta)\,\mathrm{d}\lambda$.
- In the quadratic form $X^T \sum_n^{-1} X$, we replace $\sum_n^{-1}$ by the matrix A defined as

$$A(\theta) = [\alpha(j-l)]_{j,l=1,\ldots,n} \tag{42}$$

  where

$$\alpha(j-l) = (2\pi)^{-2}\int_{-\pi}^{\pi}\frac{1}{f(\lambda,\theta)}e^{i(j-l)\lambda}\,\mathrm{d}\lambda. \tag{43}$$

- By using the two steps above, we obtain the approximate maximum log likelihood

$$L_n^* = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\frac{1}{2\pi}\int_{-\pi}^{\pi}\log f(\lambda,\theta)\,\mathrm{d}\lambda - \frac{1}{2}X^T A(\theta)X \tag{44}$$

- The Whittle estimator is given by

$$\hat{\theta} = \arg\min_{\theta} L(\theta) \tag{45}$$

where

$$L(\theta) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\lambda, \theta)\, d\lambda + \frac{X^T A(\theta) X}{n} \tag{46}$$

Fox & Taqqu (1986) and Dahlhaus (1989) showed the consistency, the asymptotic normality and the efficiency of this estimator.

Simulation Study

We carry out an experiment of 1000 samples for three processes: an ARFIMA $(1, d, 1)$, an ARFIMA $(0, d, 0)$ and a fractional Gaussian noise.[7] For the two latter processes, we use four different sample sizes ($n = 30$, $n = 200$, $n = 1000$ and $n = 10000$), whereas for the first one, we use only two sample sizes ($n = 30$ and $n = 1000$). In this design, we take d in the interval $]-0.5, 0.45]$ with a step of 0.1 for $d \in]-0.5, 0[$ and 0.05 for $d \in]0, 0.5[$, to examine the persistence and the anti-persistence effect.[8] To show the efficiency of the estimator, we determine the mean and the estimated standard deviation of $\hat{d}$ for each value of d. To check the asymptotic distribution of $\hat{d}$, we determine the histograms of 1000 replications.[9]

The Analysis of the Simulation Results for an ARFIMA $(0, d, 0)$ *and an ARFIMA* $(1, d, 1)$

The data generating processes are

$$(1 - B)^d X_t = \varepsilon_t, \text{ where } \varepsilon_t \sim \text{i.i.d.}\left(0, \sigma_\varepsilon^2\right), \quad \sigma_\varepsilon^2 = 1 \tag{47}$$

and

$$(1 - \phi_1 B)(1 - B)^d X_t = (1 - \theta_1 B)\, \varepsilon_t, \text{ where } \varepsilon_t \sim \text{i.i.d.}\left(0, \sigma_\varepsilon^2\right), \quad \sigma_\varepsilon^2 = 1 \tag{48}$$

Table 1. Hurst: fractional Gaussian noise

	$n = 30$		$n = 200$		$n = 1000$		$n = 10000$	
d	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$
−0.400	−0.221	0.212	−0.294	0.159	−0.320	0.137	−0.349	0.113
−0.300	−0.146	0.236	−0.220	0.183	−0.241	0.162	−0.270	0.141
−0.200	−0.075	0.258	−0.143	0.211	−0.158	0.189	−0.186	0.176
−0.100	−0.006	0.276	−0.064	0.241	−0.073	0.217	−0.097	0.212
0.000	0.060	0.291	0.015	0.270	0.014	0.247	−0.005	0.248
0.100	0.125	0.302	0.096	0.296	0.103	0.276	0.087	0.283
0.150	0.156	0.306	0.137	0.309	0.148	0.290	0.135	0.299
0.200	0.186	0.309	0.176	0.319	0.192	0.303	0.182	0.315
0.250	0.216	0.311	0.215	0.328	0.236	0.316	0.228	0.331
0.300	0.244	0.312	0.252	0.335	0.277	0.327	0.273	0.345
0.350	0.270	0.313	0.288	0.340	0.316	0.336	0.318	0.357
0.400	0.295	0.314	0.320	0.343	0.352	0.341	0.355	0.365
0.450	0.317	0.313	0.350	0.343	0.384	0.344	0.390	0.369
0.490	0.334	0.312	0.372	0.341	0.407	0.343	0.413	0.369

Table 2. Hurst: $(1 - B)^d X_t = \varepsilon_t$, with $\varepsilon_t \sim i.i.d.\ (0, 1)$

	$n = 30$		$n = 200$		$n = 1000$		$n = 10000$	
d	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$
−0.400	−0.143	0.230	−0.239	0.176	−0.285	0.149	−0.323	0.129
−0.300	−0.101	0.245	−0.190	0.194	−0.227	0.169	−0.252	0.152
−0.200	−0.050	0.261	−0.133	0.216	−0.159	0.193	−0.171	0.183
−0.100	−0.002	0.273	−0.065	0.242	−0.084	0.223	−0.086	0.218
0.000	0.052	0.282	0.010	0.269	−0.002	0.253	0.003	0.252
0.100	0.110	0.291	0.089	0.293	0.083	0.279	0.098	0.287
0.150	0.140	0.297	0.130	0.303	0.127	0.292	0.144	0.304
0.200	0.170	0.304	0.170	0.313	0.170	0.305	0.189	0.321
0.250	0.198	0.309	0.209	0.322	0.210	0.317	0.233	0.338
0.300	0.228	0.313	0.246	0.329	0.250	0.328	0.275	0.355
0.350	0.257	0.314	0.281	0.335	0.288	0.336	0.315	0.369
0.400	0.286	0.315	0.315	0.340	0.321	0.344	0.350	0.379
0.450	0.313	0.318	0.348	0.342	0.355	0.346	0.378	0.388
0.490	0.333	0.317	0.376	0.341	0.381	0.344	0.395	0.388

θ_1 is fixed to 0.25, whereas the parameter ϕ_1 varies from -0.8 to 0.8 with a step of 0.2.[10] For the semi-parametric methods, the truncation parameter m takes the value $n^{1/2}$.

- The results of the Hurst method are gathered in Tables 2, 3 and 4. They show the impact of the sample size n on this method (when n increases, the results improve). For the ARFIMA $(0, d, 0)$ and for $d < -0.1$, there is a wide difference between the true value of d and the average of $\hat{d}$, whereas, for $d \geq -0.1$, there is a slight improvement of the results, but in both cases, the estimators obtained are not convergent. The effects of anti-persistence and persistence are respectively checked for $d < 0$ and $d > 0$ and this for all sample sizes. For the ARFIMA $(1, d, 1)$ process, we see the impact of the parameter ϕ_1 on the estimators. The histograms[11] for the ARFIMA $(0, d, 0)$ with $n = 30$, $n = 1000$ and $d = 0.4$ (see Figures 1 and 2), show that the estimator is skewed downwards, this is obvious for $n = 1000$. We conclude that the Hurst method underestimates the parameters.
- For the Higuchi method (Tables 6, 7 and 8), for $n = 30$, the results are very bad. For the ARFIMA $(0, d, 0)$ process, we observe a significant improvement when n becomes large. In the ARFIMA $(1, d, 1)$ even if $\phi_1 = 0$, the results still remain bad, consequently the AR component can also be a source of the bias of the Higuchi estimator. The histograms (Figures 3 and 4) show, as previously, that the Higuchi method underestimates the parameters.
- Concerning the GPH method, we see in Table 10, that the averages of $\hat{d}$ are closer to the true values, moreover the estimated standard deviations are weak, whereas in Tables 11 and 12 we observe that the results become bad when the parameter $\phi \to 1$. For example for $d = 0.45$ and $d = 0.49$, their averages overtake the stationary case. The asymptotic normality is checked on the histograms (see Figures 5 and 6).
- With regard to the Robinson (1995a) method, Table 14 shows the consistency of the estimators in ARFIMA $(0, d, 0)$ process for $n = 1000$ and 10000. In ARFIMA $(1, d, 1)$ process, the results improve when the AR component becomes very small. The histograms (Figures 7 and 8) allow us to check the asymptotic normality. This method being an extension of the GPH one, it gives, as expected, best results for a large sample size n.

Table 3. Hurst: $(1-\phi_1 B)(1-B)^d X_t = (1-0.25B)\varepsilon_t$, for $n = 30$ and with $\varepsilon_t \sim i.i.d.(0, 1)$

ϕ	d	−0.4	−0.3	−0.2	−0.1	0.0	0.1	0.15	0.2	0.2	0.3	0.35	0.4	0.45	0.49
−0.8	$\hat{d}$	−4.99	−4.99	−4.98	−4.97	−2.59	−2.23	−2.02	−1.77	−1.49	−1.19	0.52	0.93	1.34	1.45
	$\hat{S}_d$	0.16	0.13	0.16	0.25	2.23	2.43	2.55	2.67	2.80	2.91	3.14	3.24	3.36	3.50
−0.6	$\hat{d}$	−4.98	−4.97	−4.95	−4.90	−1.91	−1.44	−1.15	−0.85	−0.52	−0.16	1.40	1.79	2.18	2.36
	$\hat{S}_d$	0.13	0.18	0.26	0.39	2.44	2.64	2.74	2.83	2.92	3.00	3.16	3.25	3.34	3.46
−0.4	$\hat{d}$	−4.96	−4.92	−4.84	−4.66	−1.34	−0.79	−0.48	−0.16	0.18	0.55	1.90	2.27	2.63	2.85
	$\hat{S}_d$	0.23	0.34	0.50	0.72	2.49	2.78	2.86	2.94	3.01	3.09	3.16	3.22	3.29	3.37
−0.2	$\hat{d}$	−4.68	−4.72	−4.47	−4.07	−0.79	−0.22	0.09	0.44	0.79	1.15	2.27	2.60	2.92	3.14
	$\hat{S}_d$	0.46	0.66	0.90	1.16	2.73	2.89	2.95	3.00	3.04	3.09	3.15	3.19	3.25	3.29
0.0	$\hat{d}$	−4.50	−4.13	−3.61	−2.98	−0.25	1.10	0.66	0.99	1.32	1.66	2.57	2.86	3.13	3.33
	$\hat{S}_d$	0.88	1.13	1.36	1.57	2.84	2.91	3.02	3.06	3.11	3.14	3.14	3.15	3.18	3.17
0.2	$\hat{d}$	−1.53	−1.13	−0.68	−0.18	0.36	0.94	1.25	1.55	1.85	2.15	2.42	2.69	2.98	3.21
	$\hat{S}_d$	2.39	2.52	2.68	2.83	2.94	3.05	3.09	3.13	3.16	3.18	3.18	3.19	3.20	3.23
0.4	$\hat{d}$	−2.11	−1.35	−0.54	0.29	1.10	1.64	1.91	2.17	2.43	2.68	3.19	3.39	3.60	3.66
	$\hat{S}_d$	1.74	1.81	1.85	1.87	3.06	3.13	3.16	3.17	3.16	3.13	3.03	3.00	2.97	3.02
0.6	$\hat{d}$	−0.34	0.48	1.32	2.15	2.05	2.51	2.73	2.95	3.16	3.35	3.62	3.79	3.91	3.73
	$\hat{S}_d$	1.87	1.88	1.86	1.80	3.17	3.17	3.14	3.09	3.07	3.03	2.85	2.83	2.79	2.83
0.8	$\hat{d}$	1.50	2.33	3.11	3.78	3.32	3.64	3.79	3.92	4.07	4.19	4.25	4.29	4.30	4.09
	$\hat{S}_d$	1.86	1.79	1.65	1.42	3.14	3.04	2.97	2.90	2.83	2.78	2.64	2.67	2.52	2.24

Note: The means and the standard deviations are the reported values multiplied by 10^{-1}.

Table 4. Hurst: $(1 - \phi_1 B)(1 - B)^d X_t = (1 - 0.25B)\varepsilon_t$, for $n = 1000$ and with $\varepsilon_t \sim i.i.d.(0, 1)$

ϕ	d	−0.40	−0.30	−0.20	−0.10	0.0	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.49
−0.8	$\hat{d}$	−3.94	−3.55	−2.87	−1.23	−0.91	0.13	0.66	1.19	1.70	2.21	2.70	3.16	3.64	4.04
	$\hat{S}_d$	1.20	1.30	1.52	2.06	2.22	2.59	2.76	2.94	3.11	3.28	3.44	3.62	3.77	3.86
−0.6	$\hat{d}$	−3.70	−3.16	−2.40	−2.28	−0.55	0.41	0.91	1.40	1.88	2.36	2.81	3.24	3.69	4.05
	$\hat{S}_d$	1.23	1.40	1.66	2.26	2.34	2.67	2.83	2.99	3.16	3.31	3.46	3.61	3.73	3.78
−0.4	$\hat{d}$	−3.49	−2.89	−2.13	−1.80	−0.38	0.55	1.03	1.50	1.96	2.42	2.86	3.27	3.68	4.01
	$\hat{S}_d$	1.30	1.48	1.76	2.42	2.40	2.71	2.87	3.02	3.17	3.32	3.45	3.58	3.67	3.70
−0.2	$\hat{d}$	−3.30	−2.69	−1.98	−1.32	−0.25	0.65	1.12	1.57	2.01	2.46	2.88	3.26	3.65	3.96
	$\hat{S}_d$	1.35	1.57	1.84	2.57	2.44	2.47	2.89	3.03	3.18	3.32	3.44	3.55	3.61	3.62
0.0	$\hat{d}$	−3.12	−2.51	−1.79	−0.79	−0.15	0.73	1.19	1.63	2.06	2.48	2.89	3.25	3.61	3.90
	$\hat{S}_d$	1.43	1.65	1.90	2.69	2.48	2.77	2.90	3.04	3.18	3.31	3.41	3.50	3.54	3.54
0.2	$\hat{d}$	−2.93	−2.34	−1.64	−0.87	−0.05	0.81	1.25	1.68	2.09	2.50	2.88	3.22	3.57	3.83
	$\hat{S}_d$	1.51	1.71	1.94	2.22	2.52	2.79	2.92	3.05	3.18	3.29	3.38	3.45	3.48	3.46
0.4	$\hat{d}$	−2.74	−2.17	−1.49	0.55	0.05	0.90	1.32	1.73	2.12	2.51	2.87	3.19	3.51	3.76
	$\hat{S}_d$	1.59	1.78	2.00	2.98	2.56	2.81	2.93	3.05	3.17	3.26	3.33	3.39	3.40	3.37
0.6	$\hat{d}$	−2.51	−1.94	−1.29	1.58	0.20	1.01	1.41	1.79	2.16	2.52	2.85	3.14	3.44	3.68
	$\hat{S}_d$	1.69	1.86	2.09	3.12	2.62	2.85	2.95	3.05	3.14	3.22	3.27	3.31	3.31	3.28
0.8	$\hat{d}$	−2.08	−1.53	−0.91	2.97	0.48	1.23	1.59	1.92	2.24	2.55	2.84	3.09	3.36	3.58
	$\hat{S}_d$	1.86	2.04	2.27	3.20	2.72	2.89	2.97	3.04	3.10	3.15	3.17	3.20	3.19	3.15

Note: The means and the standard deviations are the reported values multiplied by 10^{-1}.

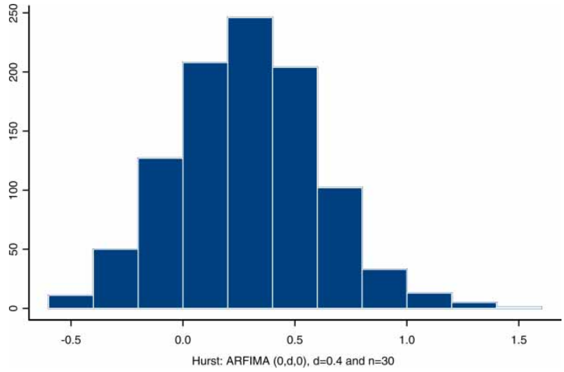

Figure 1. Hurst for $n = 30$

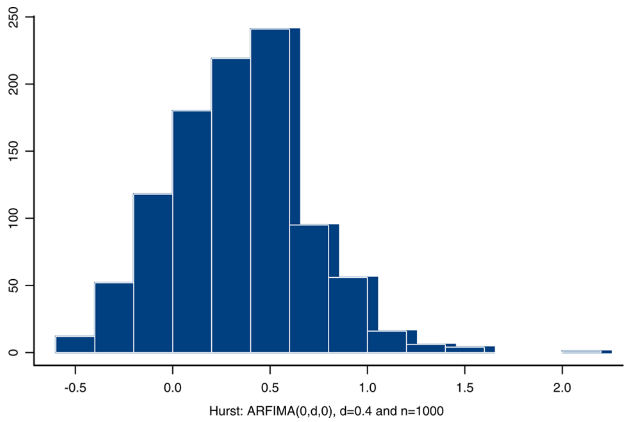

Figure 2. Hurst for $n = 1000$

Table 5. Higuchi: fractional Gaussian noise

	$n = 30$		$n = 200$		$n = 1000$		$n = 10000$	
d	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$
−0.400	−2.022	0.332	−0.503	0.065	−0.458	0.035	−0.426	0.010
−0.300	−1.974	0.371	−0.428	0.101	−0.366	0.053	−0.332	0.016
−0.200	−1.923	0.410	−0.333	0.136	−0.259	0.067	−0.223	0.020
−0.100	−1.870	0.437	−0.224	0.169	−0.146	0.076	−0.108	0.022
0.000	−1.820	0.452	−0.105	0.196	−0.023	0.082	0.006	0.024
0.100	−1.761	0.436	0.016	0.213	0.087	0.086	0.121	0.026
0.150	−1.726	0.407	0.077	0.217	0.145	0.088	0.179	0.029
0.200	−1.691	0.372	0.136	0.215	0.203	0.090	0.230	0.032
0.250	−1.660	0.348	0.195	0.210	0.260	0.090	0.294	0.037
0.300	−1.632	0.328	0.253	0.200	0.314	0.087	0.349	0.040
0.350	−1.606	0.289	0.311	0.183	0.365	0.080	0.402	0.039
0.400	−1.589	0.246	0.369	0.154	0.413	0.065	0.449	0.033
0.450	−1.624	0.384	0.412	0.132	0.443	0.059	0.478	0.031
0.490	−1.659	0.463	0.393	0.175	0.411	0.138	0.450	0.098

Table 6. Higuchi: $(1 - B)^d X_t = \varepsilon_t$, with $\varepsilon_t \sim i.i.d.(0, 1)$

	$n = 30$		$n = 200$		$n = 1000$		$n = 10000$	
d	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$
−0.400	−2.415	0.533	−0.375	0.106	−0.357	0.045	−0.337	0.014
−0.300	−2.383	0.584	−0.308	0.128	−0.286	0.055	−0.265	0.017
−0.200	−2.341	0.624	−0.230	0.149	−0.203	0.062	−0.182	0.019
−0.100	−2.282	0.648	−0.144	0.166	−0.112	0.068	−0.090	0.021
0.000	−2.209	0.653	−0.055	0.179	−0.016	0.072	0.005	0.022
0.100	−2.118	0.647	0.036	0.187	0.080	0.075	0.103	0.023
0.150	−2.073	0.654	0.081	0.189	0.128	0.076	0.152	0.024
0.200	−2.020	0.634	0.127	0.188	0.175	0.077	0.201	0.024
0.250	−1.977	0.637	0.171	0.186	0.220	0.078	0.248	0.024
0.300	−1.936	0.639	0.213	0.181	0.263	0.077	0.293	0.025
0.350	−1.890	0.614	0.252	0.176	0.302	0.075	0.334	0.024
0.400	−1.836	0.569	0.288	0.170	0.337	0.071	0.371	0.023
0.450	−1.780	0.525	0.320	0.162	0.368	0.066	0.403	0.021
0.490	−1.744	0.504	0.343	0.156	0.389	0.061	0.424	0.019

- The results of the Robinson (1995b) method, show that for $n \geq 200$, the estimators are close to the true values, and are consistent. Within this framework, the best results are obtained for small values of ϕ_1, and when $\phi_1 \to 1$, the results break down. The asymptotic normality is obtained for a large n, i.e. $n = 1000$.
- Note that, for the Whittle method (see Tables 21, 22 and 23) we have good properties in the large sample size. On the other hand for an ARFIMA $(1, d, 1)$, the results are acceptable only for $0 \leq \phi_1 \leq 0.4$. Finally, the asymptotic normality is checked only for a large n (see Figure 12).

The Whittle method is always better than the Higuchi one, and it dominates the others in large sample sizes. Obviously for an ARFIMA $(1, d, 1)$, the ϕ_1 parameter has a great influence on the results.

To appreciate the precision of an estimator, we calculate the mean squared error (MSE) given by

$$\text{MSE} = E((\hat{d} - d_0)^2) = (E(\hat{d}) - d_0)^2 + \hat{S}_d^2 \tag{49}$$

We see from Figures 13, 14, 15 and 16, that the Whittle method is the best, since it has the smallest MSE.[12]

The Analysis of the Simulation Results for FGN

For a fractional Gaussian noise, we only provide the mean squared errors for the four sample sizes and the interpretation of the results are the same as supra.[13]

Except for the Hurst method, the others give similar results in a large sample size. Concerning the Whittle method, we cannot provide the estimates; indeed, the formulas (18) and (46) show that the approximate log likelihood is a very complicated function of H, which implies important numerical difficulties.

Application to US Inflation

Generally, the Consumer Price Index P_t is modelled by an $I(1)$ process and then the inflation rate $\bar{w}_t = \log P_t - \log P_{t-1}$, an important leading indicator of the economy, is represented

Table 7. Higuchi: $(1 - \phi_1 B)(1 - B)^d X_t = (1 - 0.25B)\varepsilon_t$, for $n = 30$ and with $\varepsilon_t \sim i.i.d.(0, 1)$

ϕ	d	−0.40	−0.30	−0.20	−0.10	0.00	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.49
−0.8	$\hat{d}$	−19.7	−19.7	−19.7	−19.7	−19.7	−19.7	−19.6	−19.6	−19.6	−19.6	−19.5	−19.5	−19.4	−19.4
	$\hat{S}_d$	6.44	6.47	6.49	6.52	6.54	6.56	6.56	6.56	6.55	6.53	6.50	6.46	6.42	6.37
−0.6	$\hat{d}$	−19.7	−19.7	−19.7	−19.7	−19.7	−19.7	−19.7	−19.6	−19.6	−19.6	−19.5	−19.5	−19.4	−19.4
	$\hat{S}_d$	6.41	6.44	6.47	6.50	6.53	6.55	6.56	6.56	6.56	6.54	6.51	6.47	6.43	6.38
−0.4	$\hat{d}$	−19.7	−19.7	−19.7	−19.7	−19.7	−19.7	−19.7	−19.6	−19.6	−19.6	−19.5	−19.5	−19.4	−19.4
	$\hat{S}_d$	6.37	6.40	6.43	6.46	6.50	6.53	6.55	6.56	6.56	6.55	6.53	6.49	6.45	6.40
−0.2	$\hat{d}$	−19.7	−19.7	−19.7	−19.7	−19.7	−19.7	−19.7	−19.7	−19.6	−19.6	−19.6	−19.5	−19.5	−19.4
	$\hat{S}_d$	6.29	6.33	6.37	6.41	6.45	6.49	6.50	6.52	6.53	6.54	6.53	6.51	6.47	6.42
0.0	$\hat{d}$	−19.8	−19.8	−19.8	−19.7	−19.7	−19.7	−19.7	−19.7	−19.6	−19.6	−19.6	−19.6	−19.5	−19.5
	$\hat{S}_d$	6.16	6.22	6.27	6.32	6.37	6.42	6.45	6.47	6.49	6.50	6.51	6.51	6.49	6.45
0.2	$\hat{d}$	−19.9	−19.9	−19.8	−19.8	−19.8	−19.8	−19.7	−19.7	−19.7	−19.7	−19.6	−19.6	−19.6	−19.5
	$\hat{S}_d$	5.89	6.00	6.09	6.17	6.24	6.30	6.34	6.37	6.40	6.43	6.45	6.66	6.47	6.45
0.4	$\hat{d}$	−20.2	−20.1	−20.0	−20.0	−19.9	−19.9	−19.9	−19.9	−19.8	−19.8	−19.8	−19.7	−19.7	−19.6
	$\hat{S}_d$	5.51	5.63	5.75	5.88	5.99	6.11	6.16	6.21	6.25	6.30	6.33	6.37	6.40	6.41
0.6	$\hat{d}$	−20.9	−20.8	−20.7	−20.5	−20.4	−20.3	−20.3	−20.2	−20.2	−20.1	−20.1	−20.1	−20.4	−19.9
	$\hat{S}_d$	5.05	5.17	5.31	5.44	5.57	5.71	5.79	5.87	5.94	6.02	6.10	6.19	6.26	6.30
0.8	$\hat{d}$	−22.8	−22.6	−22.4	−22.2	−22.1	−21.9	−21.8	−21.8	−21.7	−21.6	−21.5	−21.4	−21.3	−21.2
	$\hat{S}_d$	4.45	4.56	4.69	4.85	5.04	5.25	5.36	5.46	5.57	5.67	5.79	5.90	6.02	6.11

Note: The means and the standard deviations are the reported values multiplied by 10^{-1}.

Table 8. Higuchi: $(1-\phi_1 B)(1-B)^d X_t = (1-0.25B)\varepsilon_t$, for $n = 1000$ and with $\varepsilon_t \sim i.i.d.(0, 1)$

ϕ	d	−0.40	−0.30	−0.20	−0.10	0.00	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.49
−0.8	$\hat{d}$	2.22	2.22	2.23	2.24	2.25	2.26	2.27	2.28	2.29	2.30	2.32	2.33	2.36	2.38
	$\hat{S}_d$	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.76	0.76	0.76	0.76	0.75	0.75
−0.6	$\hat{d}$	2.21	2.22	2.23	2.23	2.24	2.26	2.27	2.27	2.29	2.30	2.31	2.33	2.35	2.38
	$\hat{S}_d$	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.76	0.76	0.76	0.76	0.76	0.75	0.75
−0.4	$\hat{d}$	2.20	2.21	2.22	2.23	2.24	2.25	2.26	2.27	2.28	2.29	2.31	2.33	2.35	2.37
	$\hat{S}_d$	0.78	0.77	0.77	0.77	0.77	0.77	0.77	0.77	0.76	0.76	0.76	0.76	0.75	0.75
−0.2	$\hat{d}$	2.18	2.19	2.20	2.22	2.23	2.24	2.26	2.26	2.27	2.29	2.30	2.32	2.35	2.37
	$\hat{S}_d$	0.78	0.78	0.77	0.77	0.77	0.77	0.77	0.77	0.76	0.76	0.76	0.76	0.75	0.75
0.0	$\hat{d}$	2.14	2.15	2.17	2.19	2.20	2.22	2.23	2.24	2.26	2.27	2.29	2.31	2.33	2.36
	$\hat{S}_d$	0.78	0.78	0.78	0.78	0.78	0.78	0.77	0.77	0.77	0.76	0.76	0.76	0.76	0.75
0.2	$\hat{d}$	2.03	2.07	2.10	2.12	2.15	2.17	2.19	2.20	2.22	2.24	2.26	2.28	2.31	2.33
	$\hat{S}_d$	0.79	0.79	0.78	0.78	0.78	0.78	0.78	0.77	0.77	0.77	0.77	0.76	0.76	0.76
0.4	$\hat{d}$	1.79	1.86	1.92	1.97	2.02	2.02	2.08	2.10	2.13	2.15	2.18	2.20	2.24	2.27
	$\hat{S}_d$	0.81	0.81	0.80	0.80	0.79	0.79	0.79	0.79	0.78	0.78	0.78	0.77	0.77	0.77
0.6	$\hat{d}$	0.12	1.28	1.41	1.52	1.62	1.62	1.75	1.80	1.84	1.88	1.92	1.97	2.02	2.06
	$\hat{S}_d$	0.87	0.86	0.85	0.84	0.83	0.83	0.82	0.82	0.82	0.81	0.81	0.80	0.80	0.79
0.8	$\hat{d}$	−1.02	−0.71	−0.44	−0.19	0.03	0.03	0.35	0.45	0.55	0.66	0.76	0.86	0.97	1.06
	$\hat{S}_d$	0.87	0.88	0.90	0.90	0.09	0.90	0.90	0.90	0.90	0.90	0.90	0.89	0.89	0.89

Note: The means and the standard deviations are the reported values multiplied by 10^{-1}.

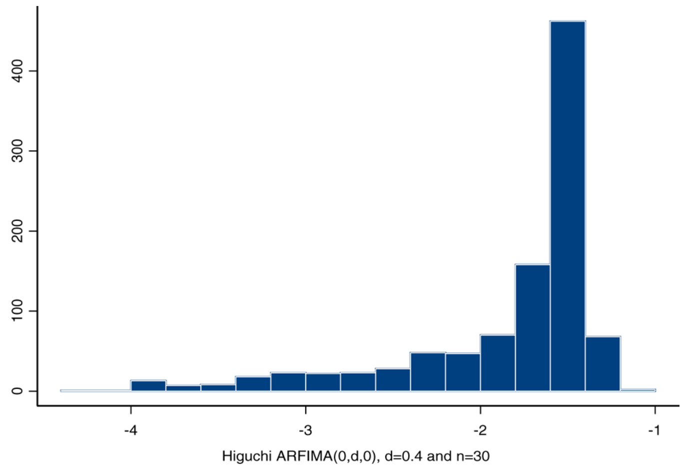

Figure 3. Higuchi for $n = 30$

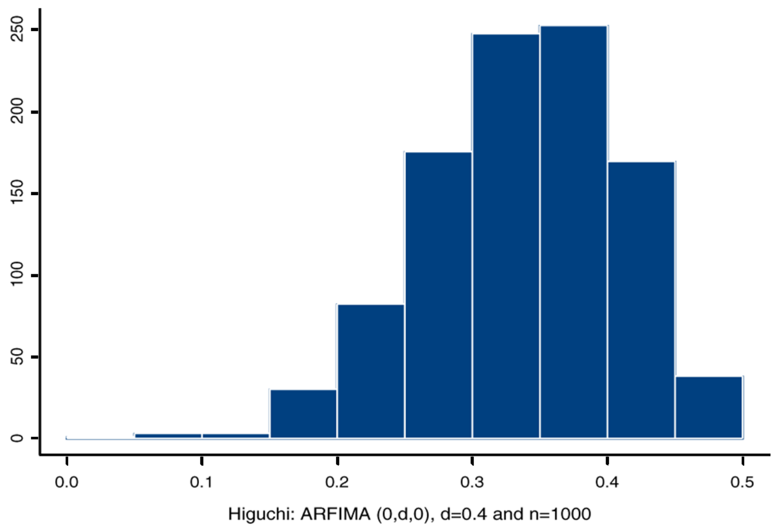

Figure 4. Higuchi for $n = 1000$

by an $I(0)$ process. But there is a controversy in the literature in considering $\bar{w}_t$ as an $I(1)$. Hassler & Wolters (1995) found that the existing tests of unit root (the augmented Dickey-Fuller test (ADF) (Dickey & Fuller, 1979), the Phillips & Perron (1988) (PP) test and the KPSS (Kwiatkowski *et al.*, 1992)) test, lead to a contradictory conclusions and hence $\bar{w}_t$ is somewhere between an $I(0)$ and $I(1)$ process. Robinson (1978), Granger (1980), Cox (1989), Taqqu *et al.* (1997), and Davidson & Sibbertsen (2004) consider that an aggregation scheme with Gaussian common component yield, after normalization, a long memory process. Our intuition is that the inflation rate is an $I(d)$ stationary process. We consider the monthly seasonally adjusted US inflation rate π_t over the period 1960: 2–2003:9 (524 observations), this indicator is an aggregate of eight leading prices (food & beverages, housing, apparel & upkeep, transportation, medical care, commodities, durables and services).[14]

Table 9. GPH: fractional Gaussian noise

	$n = 30$		$n = 200$		$n = 1000$		$n = 10000$	
d	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$
−0.400	−0.541	0.524	−0.430	0.236	−0.393	0.142	−0.389	0.071
−0.300	−0.377	0.522	−0.316	0.239	−0.294	0.138	−0.296	0.069
−0.200	−0.248	0.532	−0.211	0.239	−0.197	0.136	−0.198	0.068
−0.100	−0.125	0.530	−0.108	0.238	−0.098	0.135	−0.100	0.069
0.000	0.007	0.529	0.006	0.238	0.000	0.135	0.000	0.068
0.100	0.108	0.526	0.094	0.239	0.101	0.135	0.099	0.068
0.150	0.165	0.529	0.146	0.238	0.152	0.136	0.149	0.068
0.200	0.223	0.530	0.198	0.237	0.203	0.136	0.200	0.068
0.250	0.282	0.527	0.250	0.236	0.254	0.136	0.250	0.068
0.300	0.341	0.529	0.303	0.235	0.305	0.136	0.302	0.068
0.350	0.398	0.527	0.356	0.235	0.357	0.136	0.353	0.068
0.400	0.453	0.526	0.409	0.234	0.410	0.136	0.404	0.068
0.450	0.510	0.528	0.463	0.232	0.463	0.136	0.456	0.069
0.490	0.553	0.527	0.506	0.231	0.506	0.136	0.498	0.069

Table 10. GPH: $(1 - B)^d X_t = \varepsilon_t$, with $\varepsilon_t \sim i.i.d.(0, 1)$

	$n = 30$		$n = 200$		$n = 1000$		$n = 10000$	
d	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$
−0.400	−0.377	0.521	−0.372	0.230	−0.378	0.110	−0.391	0.072
−0.300	−0.289	0.521	−0.285	0.227	−0.289	0.143	−0.297	0.070
−0.200	−0.201	0.527	−0.192	0.226	−0.192	0.140	−0.199	0.069
−0.100	−0.107	0.529	−0.096	0.228	−0.094	0.139	−0.098	0.069
0.000	−0.004	0.519	0.001	0.231	0.005	0.139	0.002	0.070
0.100	0.098	0.512	0.102	0.228	0.107	0.139	0.104	0.070
0.150	0.149	0.513	0.153	0.229	0.159	0.138	0.156	0.070
0.200	0.201	0.514	0.203	0.230	0.210	0.137	0.207	0.070
0.250	0.253	0.515	0.254	0.230	0.261	0.138	0.259	0.070
0.300	0.304	0.510	0.306	0.231	0.313	0.139	0.310	0.071
0.350	0.354	0.514	0.358	0.231	0.366	0.143	0.362	0.071
0.400	0.405	0.514	0.410	0.232	0.420	0.140	0.414	0.071
0.450	0.457	0.511	0.463	0.233	0.475	0.141	0.466	0.071
0.490	0.498	0.513	0.507	0.232	0.520	0.140	0.508	0.071

The main characteristics of a long memory process are a slow decay of the autocorrelation function and a high peak at zero frequency of the spectral density (Granger & Joyeux, 1980). These properties are considered in a stationary framework, and we will test them. For this, we use the standard tests: ADF, PP and the KPSS test. The results are given in Table 24.[15]

This table provides contradictory results, since they show evidence of stationarity (PP) and unit root (ADF and KPSS). These statistical results and the examination of Figures 22, 24 and 26, lead us to conclude that the series can be considered as an $I(d)$ process. There is a slow decay of the autocorrelation function and a peak at zero frequency of the periodogram.

Now, we estimate the long-range dependence parameter by using the various methods described supra and we obtain Table 25.

Table 11. GPH: $(1-\phi_1 B)(1-B)^d X_t = (1-0.25B)\varepsilon_t$, for $n = 30$ and with $\varepsilon_t \sim i.i.d.(0, 1)$

ϕ	d	−0.40	−0.30	−0.20	−0.10	0.00	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.49
−0.8	$\hat{d}$	−3.90	−3.45	−2.91	−2.26	−1.47	−0.65	−0.22	0.22	0.69	1.16	1.64	2.16	2.67	3.11
	$\hat{S}_d$	5.04	5.01	5.01	5.02	4.92	5.06	5.05	5.02	5.01	5.04	5.10	5.12	5.18	5.18
−0.6	$\hat{d}$	−4.46	−3.87	−3.20	−2.44	−1.60	−0.75	−0.25	0.25	0.76	1.27	1.78	2.31	2.83	3.00
	$\hat{S}_d$	4.93	5.00	5.01	5.02	4.99	5.19	5.14	5.17	5.14	5.16	5.16	5.12	5.13	5.17
−0.4	$\hat{d}$	−4.74	−4.05	−3.27	−2.41	−1.55	−0.64	−0.14	0.36	0.89	1.40	1.92	2.44	2.95	3.36
	$\hat{S}_d$	4.93	4.97	4.94	4.98	5.16	5.19	5.20	5.20	5.25	5.22	5.17	5.18	5.14	5.16
−0.2	$\hat{d}$	−4.80	−4.02	−3.17	−2.31	−1.42	−0.48	0.04	0.56	1.09	1.60	2.12	2.64	3.15	3.57
	$\hat{S}_d$	5.10	5.00	4.99	5.06	5.16	5.28	5.20	5.16	5.18	5.11	5.09	5.15	5.15	5.14
0.0	$\hat{d}$	−4.62	−3.77	−2.89	−2.01	−1.09	−0.05	0.46	0.97	1.49	2.01	2.53	3.04	3.56	4.00
	$\hat{S}_d$	5.19	5.11	5.10	5.15	5.31	5.16	5.16	5.16	5.12	5.10	5.12	5.15	5.18	5.17
0.2	$\hat{d}$	−4.01	−3.13	−2.26	−1.33	−0.31	0.72	1.21	1.74	2.26	2.78	3.28	3.78	4.31	4.72
	$\hat{S}_d$	5.22	5.19	5.25	5.28	5.20	5.13	5.16	5.13	5.18	5.11	5.14	5.17	5.13	5.16
0.4	$\hat{d}$	−2.82	−1.93	−1.01	−0.04	1.00	2.39	2.56	3.06	3.57	4.06	4.55	5.06	5.58	6.00
	$\hat{S}_d$	5.23	5.23	5.27	5.28	5.15	5.12	5.09	5.08	5.05	5.09	5.15	5.12	5.10	5.13
0.6	$\hat{d}$	−0.08	0.09	1.09	2.11	3.17	4.19	4.68	5.18	5.65	6.15	6.66	7.15	7.65	8.03
	$\hat{S}_d$	5.31	5.39	5.28	5.22	5.10	5.05	5.03	5.03	5.10	5.12	5.05	5.04	5.08	5.06
0.8	$\hat{d}$	2.23	3.28	4.33	5.32	6.35	7.34	7.83	8.30	8.76	9.20	9.63	10.0	10.3	10.6
	$\hat{S}_d$	5.25	5.18	5.12	5.16	5.07	5.09	5.07	5.04	5.04	5.03	5.08	5.10	5.11	5.09

Note: The means and the standard deviations are the reported values multiplied by 10^{-1}.

Table 12. GPH: $(1 - \phi_1 B)(1 - B)^d X_t = (1 - 0.25B)\varepsilon_t$, for $n = 1000$ and with $\varepsilon_t \sim i.i.d.(0, 1)$

ϕ	d	−0.40	−0.30	−0.20	−0.10	0.00	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.49
−0.8	$\hat{d}$	−3.19	−2.66	−1.87	−0.95	0.02	1.02	1.54	2.05	2.57	3.09	3.62	4.16	4.71	5.16
	$\hat{S}_d$	1.61	1.48	1.41	1.39	1.40	1.41	1.38	1.37	1.38	1.39	1.40	1.40	1.40	1.39
−0.6	$\hat{d}$	−3.53	−2.81	−1.93	−0.97	0.01	1.02	1.54	2.06	2.57	3.09	3.62	4.16	4.71	5.17
	$\hat{S}_d$	1.52	1.47	1.42	1.40	1.39	1.40	1.38	1.37	1.39	1.39	1.41	1.40	1.41	1.40
−0.4	$\hat{d}$	−3.65	−2.86	−1.95	−0.98	0.02	1.02	1.54	2.06	2.57	3.09	3.62	4.16	4.72	5.17
	$\hat{S}_d$	1.47	1.43	1.41	1.40	1.39	1.39	1.38	1.37	1.38	1.40	1.40	1.41	1.41	1.40
−0.2	$\hat{d}$	−3.72	−2.90	−1.95	−0.97	0.01	1.03	1.55	2.06	2.58	3.10	3.63	4.17	4.72	5.17
	$\hat{S}_d$	1.45	1.42	1.40	1.39	1.39	1.39	1.38	1.37	1.38	1.39	1.42	1.40	1.40	1.40
0.0	$\hat{d}$	−3.76	−2.89	−1.94	−0.96	0.02	1.04	1.56	2.07	2.59	3.11	3.63	4.17	4.73	5.18
	$\hat{S}_d$	1.42	1.43	1.40	1.39	1.40	1.39	1.37	1.37	1.38	1.39	1.42	1.40	1.40	1.40
0.2	$\hat{d}$	−3.78	−3.13	−1.88	−0.94	0.04	1.06	1.58	2.09	2.61	3.13	3.65	4.19	4.75	5.20
	$\hat{S}_d$	1.43	1.44	1.34	1.39	1.39	1.39	1.38	1.37	1.38	1.39	1.43	1.40	1.41	1.40
0.4	$\hat{d}$	−3.76	−2.74	−1.84	−0.90	0.09	1.11	1.62	2.14	2.65	3.17	3.69	4.24	4.79	5.24
	$\hat{S}_d$	1.43	1.42	1.39	1.39	1.39	1.39	1.37	1.37	1.38	1.39	1.43	1.40	1.40	1.40
0.6	$\hat{d}$	−3.65	−2.61	−1.70	−0.76	0.24	1.25	1.77	2.28	2.79	3.31	3.83	4.38	4.92	5.37
	$\hat{S}_d$	1.42	1.43	1.39	1.39	1.39	1.39	1.37	1.37	1.38	1.39	1.41	1.40	1.41	1.39
0.8	$\hat{d}$	−2.99	−1.94	−1.03	−0.07	0.93	1.94	2.45	2.96	3.47	3.99	4.51	3.83	5.59	6.04
	$\hat{S}_d$	1.42	1.42	1.40	1.39	1.38	1.38	1.37	1.37	1.38	1.38	1.40	1.41	1.39	1.39

Note: The means and the standard deviations are the reported values multiplied by 10^{-1}.

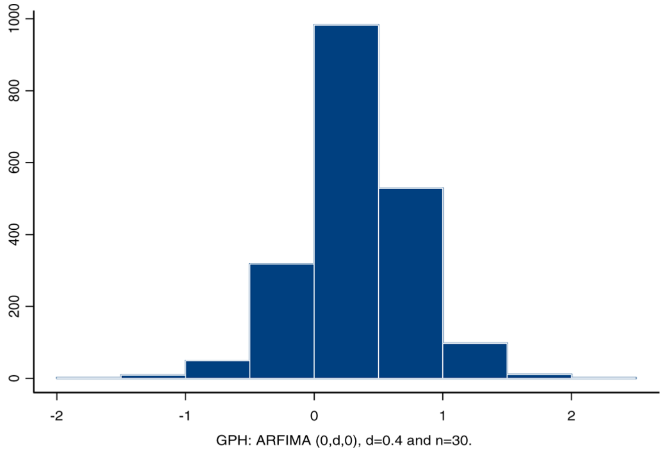

Figure 5. GPH for $n = 30$

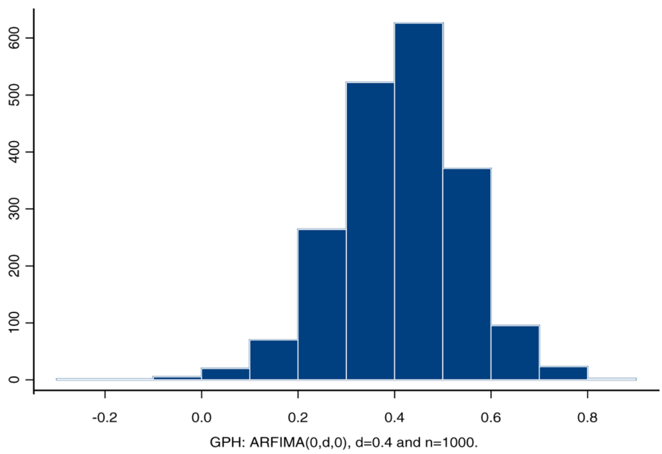

Figure 6. GPH for $n = 1000$

Table 13. Robinson (1995a): fractional Gaussian noise

	$n = 30$		$n = 200$		$n = 1000$		$n = 10000$	
d	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$
−0.400	−0.522	3.041	−0.480	0.437	−0.415	0.208	−0.397	0.082
−0.300	−0.325	3.031	−0.339	0.434	−0.304	0.208	−0.299	0.081
−0.200	−0.176	3.001	−0.226	0.433	−0.202	0.206	−0.200	0.080
−0.100	−0.044	3.011	−0.119	0.436	−0.101	0.205	−0.101	0.079
0.000	0.076	3.066	−0.016	0.433	0.000	0.204	0.000	0.079
0.100	0.197	3.133	0.086	0.430	0.099	0.204	0.098	0.079
0.150	0.268	3.119	0.137	0.429	0.149	0.204	0.149	0.079
0.200	0.334	3.133	0.188	0.428	0.200	0.204	0.200	0.079
0.250	0.401	3.159	0.239	0.429	0.251	0.204	0.249	0.079
0.300	0.472	3.178	0.291	0.429	0.302	0.204	0.300	0.079
0.350	0.541	3.194	0.344	0.428	0.354	0.204	0.351	0.079
0.400	0.610	3.221	0.396	0.425	0.406	0.203	0.402	0.079
0.450	0.689	3.317	0.448	0.424	0.458	0.203	0.453	0.079
0.490	0.723	3.266	0.490	0.424	0.500	0.203	0.495	0.080

Table 14. Robinson (1995a):$(1-B)^d X_t = \varepsilon_t$, with $\varepsilon_t \sim i.i.d.(0, 1)$

	$n = 30$		$n = 200$		$n = 1000$		$n = 10000$	
d	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$
−0.400	−0.380	1.861	−0.374	0.391	−0.387	0.187	−0.388	0.081
−0.300	−0.285	1.845	−0.279	0.385	−0.294	0.189	−0.292	0.080
−0.200	−0.193	1.846	−0.184	0.388	−0.196	0.187	−0.194	0.080
−0.100	−0.103	1.852	−0.086	0.391	−0.098	0.187	−0.095	0.080
0.000	−0.003	1.861	0.013	0.397	0.001	0.187	0.005	0.080
0.100	0.098	1.861	0.117	0.397	0.102	0.186	0.107	0.080
0.150	0.154	1.853	0.167	0.399	0.154	0.185	0.159	0.080
0.200	0.214	1.857	0.218	0.400	0.205	0.185	0.211	0.080
0.250	0.272	1.844	0.269	0.400	0.256	0.186	0.263	0.080
0.300	0.327	1.826	0.320	0.401	0.308	0.187	0.316	0.081
0.350	0.384	1.814	0.372	0.401	0.360	0.188	0.369	0.081
0.400	0.439	1.809	0.424	0.397	0.413	0.189	0.422	0.082
0.450	0.491	1.816	0.477	0.396	0.467	0.191	0.476	0.082
0.490	0.535	1.819	0.520	0.396	0.511	0.191	0.519	0.082

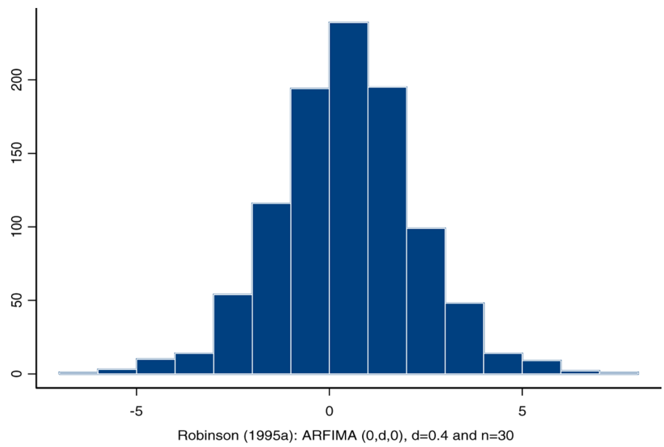

Figure 7. Robinson (1995a), $n = 30$

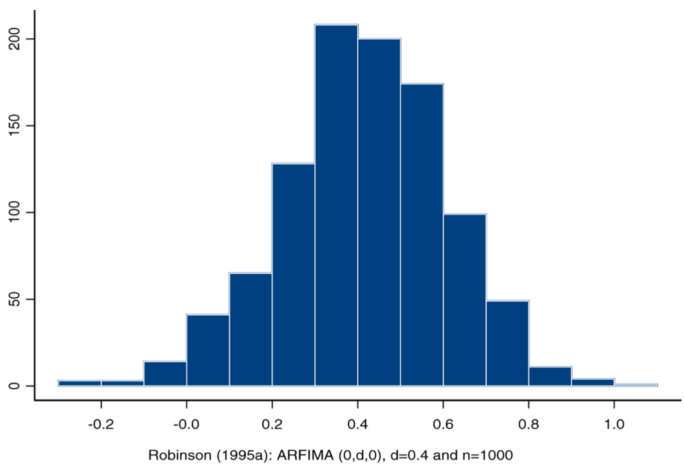

Figure 8. Robinson (1995a), $n = 1000$

Table 15. Robinson (1995a): $(1-\phi_1 B)(1-B)^d X_t = (1-0.25B)\varepsilon_t$, for $n = 30$ and with $\varepsilon_t \sim i.i.d.(0, 1)$

ϕ	d	−0.40	−0.30	−0.20	−0.10	0.00	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.49
−0.8	$\hat{d}$	−6.29	−5.75	−5.04	−4.33	−3.62	−2.77	−2.34	−1.93	−1.52	−1.02	−0.46	0.08	0.69	1.17
	$\hat{S}_d$	17.3	17.5	17.7	18.0	18.5	18.8	18.9	18.8	18.9	18.9	18.8	19.0	18.9	18.9
−0.6	$\hat{d}$	−6.99	−6.24	−5.40	−4.64	−3.91	−3.07	−2.56	−2.06	−1.58	−1.08	0.50	0.08	0.70	1.18
	$\hat{S}_d$	18.3	18.4	18.6	18.5	18.6	18.8	18.7	18.7	18.8	19.0	18.9	18.7	18.4	18.3
−0.4	$\hat{d}$	−6.99	−6.20	−5.35	−4.54	−3.72	−2.80	−2.33	−1.78	−1.19	−0.64	−0.06	0.54	1.11	1.61
	$\hat{S}_d$	18.2	18.4	18.4	18.7	18.8	19.0	19.1	19.0	19.1	18.9	18.8	18.8	18.3	18.2
−0.2	$\hat{d}$	−6.57	−5.81	−4.99	−4.15	−3.26	−2.27	−1.74	−1.19	−0.56	−0.05	0.50	1.15	1.74	2.23
	$\hat{S}_d$	18.6	18.7	18.7	18.9	18.9	18.8	18.7	18.7	18.7	18.5	18.5	18.5	18.1	18.1
0.0	$\hat{d}$	−5.83	−4.98	−4.12	−3.24	−2.26	−1.22	−0.65	1.01	0.46	1.02	1.61	2.20	2.79	3.28
	$\hat{S}_d$	18.7	18.7	18.8	18.8	18.7	18.6	18.6	18.5	18.4	18.4	18.6	18.2	18.2	18.3
0.2	$\hat{d}$	−4.33	−3.37	−2.45	−1.53	−0.56	0.46	1.01	1.58	2.20	2.74	3.30	3.88	4.43	4.85
	$\hat{S}_d$	18.7	18.5	18.5	18.5	18.5	18.6	18.5	18.5	18.5	18.3	18.1	18.1	18.1	18.1
0.4	$\hat{d}$	−1.96	−1.04	−0.12	0.84	1.76	2.87	3.45	4.03	4.59	5.12	5.64	6.15	6.70	7.16
	$\hat{S}_d$	18.4	18.4	18.5	18.5	18.7	18.7	18.6	18.4	18.2	18.1	18.0	18.1	18.1	18.3
0.6	$\hat{d}$	0.73	1.70	2.69	3.67	4.74	5.84	6.33	6.85	7.39	7.91	8.39	8.89	9.44	9.82
	$\hat{S}_d$	18.5	18.6	18.6	18.6	18.5	18.4	18.4	18.2	18.1	18.1	18.2	18.2	18.6	18.5
0.8	$\hat{d}$	3.03	4.02	5.13	6.20	7.33	8.37	8.84	9.34	9.84	10.3	10.8	11.1	1.15	1.71
	$\hat{S}_d$	18.4	18.5	18.4	18.3	18.0	18.1	18.1	18.1	18.1	17.8	18.0	18.0	17.8	17.7

Note: The means and the standard deviations are the reported values multiplied by 10^{-1}.

Table 16. Robinson (1995a): $(1 - \phi_1 B)\,(1 - B)^d X_t = (1 - 0.25B)\varepsilon_t$, for $n = 1000$ and with $\varepsilon_t \sim i.i.d.(0, 1)$

ϕ	d	−0.40	−0.30	−0.20	−0.10	0.00	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.49
−0.8	$\hat{d}$	−3.37	−2.73	−1.92	−1.01	−0.04	0.98	1.48	1.99	2.51	3.02	3.55	4.08	4.62	5.06
	$\hat{S}_d$	2.01	1.90	1.86	1.87	1.88	1.86	1.84	1.84	1.86	1.86	1.87	1.89	1.91	1.90
−0.6	$\hat{d}$	−3.67	−2.88	−1.98	−1.02	−0.03	0.96	1.48	1.99	2.50	3.02	3.55	4.08	4.62	5.06
	$\hat{S}_d$	1.92	1.89	1.88	1.88	1.87	1.86	1.85	1.85	1.86	1.86	1.88	1.88	1.91	1.91
−0.4	$\hat{d}$	−3.79	−2.93	−2.00	−1.02	−0.03	0.97	1.48	1.99	2.50	3.03	3.55	4.08	4.62	5.06
	$\hat{S}_d$	1.88	1.89	1.89	1.87	1.88	1.86	1.85	1.85	1.86	1.86	1.87	1.88	1.91	1.91
−0.2	$\hat{d}$	−3.84	−2.95	−2.01	−1.02	−0.03	0.98	1.49	2.00	2.51	3.03	3.56	4.09	4.63	5.07
	$\hat{S}_d$	1.86	1.88	1.89	1.87	1.88	1.85	1.85	1.85	1.86	1.87	1.88	1.89	1.91	1.91
0.0	$\hat{d}$	−3.87	−2.96	−1.99	−1.01	−0.03	0.99	1.50	2.01	2.52	3.04	3.57	4.08	4.64	5.08
	$\hat{S}_d$	1.87	1.89	1.88	1.87	1.88	1.85	1.85	1.85	1.86	1.87	1.88	1.89	1.91	1.91
0.2	$\hat{d}$	−3.88	−2.94	−1.97	−0.99	0.04	1.01	1.53	2.04	2.55	3.07	3.60	4.12	4.66	5.10
	$\hat{S}_d$	1.87	1.89	1.87	1.87	1.87	1.85	1.85	1.85	1.86	1.87	1.88	1.89	1.92	1.91
0.4	$\hat{d}$	−3.84	−2.89	−1.91	−0.93	0.07	1.08	1.59	2.10	2.61	3.13	3.66	4.19	4.72	5.16
	$\hat{S}_d$	1.88	1.88	1.85	1.87	1.86	1.86	1.85	1.85	1.86	1.87	1.87	1.89	1.92	1.91
0.6	$\hat{d}$	−3.67	−2.70	−1.72	−0.73	0.27	1.28	1.79	2.30	2.81	3.33	3.85	4.38	4.91	5.35
	$\hat{S}_d$	1.89	1.88	1.88	1.87	1.86	1.86	1.85	1.86	1.87	1.87	1.87	1.89	1.92	1.92
0.8	$\hat{d}$	−2.72	−1.74	−0.75	0.23	1.23	2.24	2.75	3.25	3.76	4.28	4.80	5.32	5.85	6.29
	$\hat{S}_d$	1.89	1.88	1.87	1.88	1.87	1.86	1.85	1.86	1.87	1.87	1.87	1.90	1.91	1.93

Note: The means and the standard deviations are the reported values multiplied by 10^{-1}.

Table 17. Robinson (1995b): fractional Gaussian noise

	$n = 30$		$n = 200$		$n = 1000$		$n = 10000$	
d	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$
−0.400	−0.486	1.329	−0.454	0.216	−0.399	0.122	−0.395	0.058
−0.300	−0.353	1.331	−0.338	0.213	−0.300	0.117	−0.301	0.056
−0.200	−0.246	1.349	−0.234	0.212	−0.203	0.116	−0.202	0.056
−0.100	−0.140	1.322	−0.133	0.211	−0.105	0.116	−0.103	0.056
0.000	−0.033	1.310	−0.032	0.211	−0.006	0.116	−0.005	0.055
0.100	0.078	1.306	0.069	0.211	0.093	0.116	0.094	0.055
0.150	0.133	1.312	0.120	0.210	0.144	0.116	0.145	0.055
0.200	0.189	1.311	0.172	0.210	0.195	0.116	0.197	0.056
0.250	0.254	1.300	0.224	0.210	0.246	0.117	0.246	0.056
0.300	0.311	1.297	0.277	0.209	0.298	0.117	0.298	0.056
0.350	0.370	1.295	0.330	0.209	0.350	0.117	0.348	0.056
0.400	0.430	1.309	0.383	0.209	0.402	0.117	0.399	0.056
0.450	0.488	1.326	0.436	0.208	0.454	0.118	0.451	0.056
0.490	0.531	1.336	0.479	0.208	0.497	0.118	0.492	0.057

Table 18. Robinson (1995b): $(1 - B)^d X_t = \varepsilon_t$, with $\varepsilon_t \sim i.i.d.(0, 1)$

	$n = 30$		$n = 200$		$n = 1000$		$n = 10000$	
d	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$
−0.400	−0.250	0.320	−0.349	0.145	−0.376	0.097	−0.389	0.060
−0.300	−0.203	0.329	−0.284	0.170	−0.295	0.108	−0.298	0.062
−0.200	−0.150	0.344	−0.204	0.185	−0.203	0.113	−0.202	0.062
−0.100	−0.093	0.357	−0.114	0.194	−0.105	0.113	−0.103	0.062
0.000	−0.031	0.366	−0.018	0.196	−0.007	0.114	−0.002	0.062
0.100	0.032	0.369	0.082	0.194	0.093	0.114	0.099	0.062
0.150	0.063	0.368	0.133	0.191	0.144	0.114	0.150	0.062
0.200	0.095	0.365	0.185	0.187	0.196	0.114	0.202	0.062
0.250	0.125	0.361	0.236	0.180	0.247	0.113	0.254	0.063
0.300	0.154	0.356	0.286	0.171	0.299	0.110	0.306	0.063
0.350	0.162	0.350	0.334	0.159	0.349	0.106	0.358	0.062
0.400	0.208	0.340	0.377	0.143	0.395	0.096	0.410	0.059
0.450	0.232	0.330	0.414	0.125	0.434	0.082	0.454	0.049
0.490	0.248	0.324	0.438	0.109	0.458	0.067	0.478	0.035

In order to check the adequacy of the model, we apply the diagnostic tests over the residuals. For this, we truncate the filter of equation (2) in the following way

$$(1 - B)^{\hat{d}} = \sum_{j=0}^{k} \tau_j B^j, \quad \tau_j = \frac{j - 1 - \hat{d}}{j} \tau_{j-1}, \quad \tau_0 = 1$$

the value of k is obtained by checking the following condition (Hassler & Wolters, 1995)

$$|\tau_{k-1}| \geq 0.005, \quad |\tau_k| < 0.005$$

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

$$\hat{\varepsilon}_t = \sum_{j=0}^{k} \tau_j \pi_{t-j}, \quad t = k + 1, k + 2, \ldots, n$$

Table 19. Robinson (1995b): $(1-\phi_1 B)\,(1-B)^d X_t = (1-0.25B)\varepsilon_t$, for $n = 30$ and with $\varepsilon_t \sim i.i.d.(0,1)$

ϕ	d	−0.40	−0.30	−0.20	−0.10	0.00	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.49
−0.8	$\hat{d}$	−2.25	−2.06	−1.78	−1.40	−0.95	−0.43	−0.14	0.13	0.41	0.72	1.02	1.33	1.62	1.85
	$\hat{S}_d$	3.17	3.26	3.33	3.43	3.54	3.60	3.64	3.66	3.67	3.65	3.64	3.59	3.53	3.47
−0.6	$\hat{d}$	−2.53	−2.27	−1.91	−1.48	−0.97	−0.40	−0.09	0.21	0.52	0.83	1.13	1.42	1.71	1.93
	$\hat{S}_d$	3.13	3.24	3.35	3.46	3.55	3.61	3.64	3.67	3.66	3.66	3.63	3.58	3.51	3.46
−0.4	$\hat{d}$	−2.69	−2.37	−1.97	−1.49	−0.94	−0.35	−0.04	0.27	0.59	0.89	1.19	1.48	1.71	1.98
	$\hat{S}_d$	3.06	3.18	3.32	3.45	3.55	3.64	3.66	3.67	0.67	3.65	3.63	3.58	3.51	3.44
−0.2	$\hat{d}$	−2.76	−2.39	−1.96	−1.45	−0.88	−0.28	0.03	0.35	−1.62	0.98	1.28	1.57	1.84	2.04
	$\hat{S}_d$	3.02	3.17	3.31	3.45	3.57	3.66	3.67	3.68	1.81	3.65	3.62	3.56	3.48	3.41
0.0	$\hat{d}$	−2.74	−2.33	−1.86	−1.33	−0.74	−0.12	0.19	0.51	−0.12	1.14	1.43	1.71	1.98	2.17
	$\hat{S}_d$	3.02	3.17	3.33	3.48	3.61	3.67	3.68	3.68	1.87	3.64	3.59	3.53	3.44	3.36
0.2	$\hat{d}$	−2.57	−2.12	−1.61	−1.04	−0.43	0.19	0.52	0.83	1.52	1.44	1.72	1.98	2.23	2.40
	$\hat{S}_d$	3.08	3.26	3.42	3.55	3.65	3.68	3.68	3.66	1.86	3.58	3.52	3.44	3.34	3.27
0.4	$\hat{d}$	−2.11	−1.60	−1.03	−0.42	0.19	0.82	1.13	1.43	3.17	1.99	2.24	2.47	2.67	2.80
	$\hat{S}_d$	3.27	3.42	3.55	3.64	3.68	3.66	3.62	3.58	1.62	3.45	3.35	3.24	3.15	3.10
0.6	$\hat{d}$	−1.04	−0.45	0.15	0.76	1.37	1.94	2.21	2.46	4.39	2.89	3.07	3.21	3.32	3.39
	$\hat{S}_d$	3.53	3.63	3.67	3.65	3.57	3.44	3.36	3.26	1.06	3.04	2.93	2.84	2.78	2.74
0.8	$\hat{d}$	0.69	1.56	2.13	2.64	3.07	3.43	3.58	3.71	4.87	3.90	3.95	3.97	3.98	3.96
	$\hat{S}_d$	3.62	3.54	3.39	3.17	2.95	2.71	2.58	2.46	0.47	2.29	2.28	2.30	2.32	2.33

Note: The means and the standard deviations are the reported values multiplied by 10^{-1}.

Table 20. Robinson (1995b): $(1-\phi_1 B)\,(1-B)^d X_t = (1-0.25B)\varepsilon_t$, for $n = 1000$ and with $\varepsilon_t \sim i.i.d.(0,1)$

ϕ	d	−0.40	−0.30	−0.20	−0.10	0.00	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.49
−0.8	$\hat{d}$	−3.27	−2.74	−1.98	−1.01	−0.10	0.89	1.40	1.91	2.43	2.95	3.46	3.92	4.32	4.57
	$\hat{S}_d$	1.24	1.17	1.16	1.15	1.14	1.14	1.14	1.13	1.13	1.11	1.06	0.97	0.83	0.68
−0.6	$\hat{d}$	−3.54	−2.87	−2.02	−1.08	−0.10	0.89	1.40	1.92	2.43	2.95	3.46	3.92	4.32	4.57
	$\hat{S}_d$	1.10	1.12	1.14	1.14	1.14	1.14	1.14	1.13	1.13	1.11	1.06	0.97	0.83	0.68
−0.4	$\hat{d}$	−3.65	−2.92	−2.04	−1.08	−0.10	0.89	1.40	1.92	2.43	2.95	3.46	3.93	4.32	4.57
	$\hat{S}_d$	1.04	1.10	1.13	1.14	1.14	1.14	1.14	1.13	1.13	1.11	1.06	0.97	0.83	0.68
−0.2	$\hat{d}$	−3.71	−2.94	−2.04	−1.08	−0.10	0.90	1.41	1.92	2.44	2.96	3.46	3.93	4.32	4.57
	$\hat{S}_d$	0.10	1.09	1.13	1.14	1.14	1.14	1.14	1.13	1.13	1.11	1.06	0.97	0.82	0.68
0.0	$\hat{d}$	−3.74	−2.95	−2.04	−1.08	−0.09	0.91	1.42	1.93	2.45	2.96	3.47	3.94	4.33	4.57
	$\hat{S}_d$	0.98	1.09	1.13	1.13	1.14	1.14	1.14	1.13	1.13	1.11	1.06	0.97	0.82	0.68
0.2	$\hat{d}$	−3.76	−2.95	−2.04	−1.06	−0.07	0.93	1.43	1.95	2.47	2.98	3.49	3.95	4.34	4.58
	$\hat{S}_d$	0.97	1.08	1.13	1.13	1.14	1.14	1.14	1.13	1.13	1.10	1.06	0.96	0.82	0.67
0.4	$\hat{d}$	−3.75	−2.92	−2.03	−1.02	−0.03	0.97	1.48	1.99	2.51	3.02	3.52	3.98	4.37	4.60
	$\hat{S}_d$	0.97	1.08	1.13	1.13	1.14	1.14	1.14	1.13	1.12	1.10	1.05	0.95	0.80	0.66
0.6	$\hat{d}$	−3.66	−2.80	−1.99	−0.89	0.10	1.10	1.61	2.12	2.64	3.15	3.64	4.08	4.44	4.65
	$\hat{S}_d$	0.98	1.09	1.12	1.13	1.14	1.14	1.14	1.13	1.12	1.09	1.03	0.92	0.76	0.62
0.8	$\hat{d}$	−3.08	−2.16	−1.86	−0.21	0.78	1.87	2.29	2.79	3.29	0.12	4.18	4.51	4.73	4.85
	$\hat{S}_d$	1.07	1.13	1.13	1.14	1.14	1.14	1.13	1.11	1.08	1.00	0.88	0.72	0.54	0.40

Note: The means and the standard deviations are the reported values multiplied by 10^{-1}.

Table 21. Whittle: $(1-B)^d X_t = \varepsilon_t$, with $\varepsilon_t \sim i.i.d.(0, 1)$

	$n = 30$		$n = 200$		$n = 1000$		$n = 10000$	
d	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$	$\hat{d}$	$\hat{S}_d$
−0.400	−0.441	0.102	−0.421	0.055	−0.406	0.027	−0.400	0.008
−0.300	−0.404	0.129	−0.327	0.064	−0.306	0.026	−0.300	0.007
−0.200	−0.341	0.160	−0.230	0.065	−0.205	0.027	−0.200	0.008
−0.100	−0.275	0.189	−0.131	0.068	−0.105	0.028	−0.101	0.008
0.000	−0.194	0.207	−0.030	0.067	−0.005	0.027	0.000	0.007
0.100	−0.090	0.212	0.071	0.065	0.092	0.026	0.099	0.007
0.150	−0.034	0.221	0.120	0.066	0.143	0.028	0.149	0.008
0.200	0.015	0.228	0.171	0.064	0.195	0.027	0.199	0.008
0.250	0.043	0.231	0.227	0.062	0.243	0.026	0.249	0.008
0.300	0.119	0.222	0.277	0.066	0.294	0.025	0.299	0.007
0.350	0.148	0.226	0.323	0.067	0.345	0.024	0.349	0.009
0.400	0.204	0.219	0.373	0.064	0.396	0.025	0.388	0.008
0.450	0.247	0.216	0.422	0.059	0.445	0.026	0.449	0.007
0.490	0.287	0.210	0.454	0.052	0.481	0.020	0.489	0.007

For each estimated d, we compute the corresponding vector of the residuals, and we apply the Ljung-Box (LB) portmanteau standard test and the Shapiro-Wilk (SW) normality test. The results are regrouped in Table 26.

These results show that the model is not adequate and, hence, we will fit an $AR(p)$ process by using the AIC criterion, a short memory component for $\hat{\varepsilon}_t$ (Table 27).

According to the simulation results, we observe the same hierarchy between the different methods. We conclude that the US inflation rate is generated by an ARFIMA $(1, d, 0)$ and all the methods give, except the Higuchi one, d approximately equal to 0.45 (i.e. the seasonally adjusted US inflation is an ARFIMA $(1, 0.45, 0)$ and $\hat{\phi}_1$ depends on the considered method).

The Hurst and Higuchi methods give $d = 0.532$ and $d = 0.993$ respectively, hence, the time series is non-stationary which contradicts the results of the PP test. These estimators are not reliable and we think that this problem is due to the fact that these two methods are adapted to the self-similar process not to the ARFIMA one. That is why we did not consider these methods to check the adequacy of the model.

As with Robinson (1978), Granger (1980), Cox (1989), Taqqu *et al.* (1997), and Davidson & Sibbertsen (2004), we confirm in this illustration that the aggregation procedure generates a long memory.

Conclusion

The results of the application of some estimation methods on simulated data, showed that the Hurst and the Higuchi estimators do not have good properties in terms of the consistency and the asymptotic normality, whereas they have been verified for the Geweke & Porter-Hudak (1983), Robinson (1995a, b) as well as the Whittle estimators. So we can conclude that there is a commensurability between the theoretical and empirical results for the semi-parametric and the maximum likelihood methods, especially for a large sample size ($n = 1000$). The comparison of the different estimators showed that the Whittle estimator is the best one, since it has the small mean squared error. These conclusions are true only for the ARFIMA $(0, d, 0)$ process, whereas they vary with the AR component ϕ_1, which has an important

Table 22. Whittle: $(1-\phi_1 B)\,(1-B)^d X_t = (1-0.25B)\varepsilon_t$, for $n = 30$ and with $\varepsilon_t \sim i.i.d.(0, 1)$

ϕ	d	−0.40	−0.30	−0.20	−0.10	0.00	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.49
−0.8	$\hat{d}$	−4.99	−4.99	−4.99	−4.99	−4.99	−4.98	−4.98	−4.95	−4.94	−4.90	−4.83	−4.73	−4.65	−4.50
	$\hat{S}_d$	0.00	0.00	0.00	0.00	0.13	0.09	0.17	0.23	0.29	0.40	0.58	0.71	0.84	1.04
−0.6	$\hat{d}$	−4.99	−4.99	−4.99	−4.99	−4.97	−4.92	−4.87	−4.80	−4.66	−4.53	−4.30	−4.12	−3.88	−3.53
	$\hat{S}_d$	0.00	0.01	0.02	0.09	0.19	0.32	0.44	0.56	0.77	0.96	1.13	1.26	1.40	1.65
−0.4	$\hat{d}$	−4.98	−4.99	−4.98	−4.93	−4.85	−4.66	−4.51	−4.31	−4.15	−3.77	−3.47	−3.07	−2.57	−2.21
	$\hat{S}_d$	0.03	0.05	0.10	0.32	0.52	0.75	0.94	1.10	1.23	1.50	1.58	1.83	1.93	2.01
−0.2	$\hat{d}$	−4.98	−4.97	−4.89	−4.79	−4.46	−4.09	−3.80	−3.41	−3.04	−2.61	−2.16	−1.71	−1.29	−0.80
	$\hat{S}_d$	0.14	0.15	0.40	0.59	0.97	1.25	1.45	1.54	1.88	1.94	2.06	2.14	2.22	2.24
0.0	$\hat{d}$	−4.92	−4.79	−4.58	−4.19	−3.70	−2.89	−2.49	−2.11	−1.80	−1.24	−0.63	−0.29	0.22	0.68
	$\hat{S}_d$	0.32	0.60	0.85	1.17	1.47	1.78	1.89	1.99	2.09	2.18	2.22	2.24	2.37	2.28
0.2	$\hat{d}$	−4.62	−4.31	−3.76	−3.06	−2.23	−1.32	−0.83	−0.32	−0.02	0.59	1.65	1.70	2.07	2.39
	$\hat{S}_d$	0.84	1.10	1.42	1.70	1.92	2.13	2.14	2.15	2.28	2.29	2.20	2.26	2.15	2.30
0.4	$\hat{d}$	−3.59	−2.95	−2.27	−1.30	−0.45	0.55	1.12	1.55	2.01	2.53	2.97	3.42	3.74	3.92
	$\hat{S}_d$	1.49	1.78	1.95	2.13	2.18	2.21	2.20	2.28	2.30	2.09	1.90	1.86	1.68	1.60
0.6	$\hat{d}$	−1.93	−0.86	−0.09	0.73	1.78	2.63	3.00	3.36	3.79	4.11	4.23	4.45	4.60	4.74
	$\hat{S}_d$	2.03	2.24	2.19	2.23	2.17	2.10	1.91	1.89	1.62	1.43	1.35	1.21	0.96	0.85
0.8	$\hat{d}$	0.15	1.21	2.12	2.89	3.59	4.14	4.39	4.56	4.64	4.78	4.87	4.90	4.95	4.98
	$\hat{S}_d$	2.32	2.23	2.17	2.02	1.76	1.47	1.23	1.06	1.00	0.71	0.49	0.49	0.29	0.23

Note: The means and the standard deviations are the reported values multiplied by 10^{-1}.

Table 23. Whittle: $(1 - \phi_1 B)\,(1 - B)^d X_t = (1 - 0.25B)\varepsilon_t$, for $n = 1000$ and with $\varepsilon_t \sim i.i.d.(0, 1)$

ϕ	d	−0.40	−0.30	−0.20	−0.10	0.00	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.49
−0.8	$\hat{d}$	−4.99	−4.99	−4.99	−4.99	−4.98	−4.52	−4.07	−3.61	−3.09	−2.57	−2.05	−1.56	−1.03	−0.62
	$\hat{S}_d$	0.00	0.00	0.00	0.00	0.07	0.38	0.46	0.47	0.48	0.48	0.47	0.47	0.48	0.47
−0.6	$\hat{d}$	−4.99	−4.99	−4.99	−4.99	−4.43	−3.47	−2.79	−2.46	−1.98	−1.44	−0.94	−0.46	0.05	0.49
	$\hat{S}_d$	0.00	0.00	0.00	0.09	0.38	0.40	0.41	0.40	0.43	0.42	0.41	0.40	0.40	0.41
−0.4	$\hat{d}$	−4.99	−4.99	−4.99	−4.93	−3.55	−2.56	−2.05	−1.57	−1.07	−0.55	−0.31	0.45	0.95	1.38
	$\hat{S}_d$	0.00	0.00	0.03	0.32	0.34	0.34	0.37	0.36	0.37	0.38	0.37	0.37	0.37	0.37
−0.2	$\hat{d}$	−4.99	−4.97	−4.64	−4.79	−2.64	−1.66	−1.77	−0.65	−0.15	0.34	0.85	1.37	1.86	2.28
	$\hat{S}_d$	0.00	0.01	0.26	0.59	0.32	0.30	0.32	0.33	0.31	0.31	0.32	0.32	0.31	0.33
0.0	$\hat{d}$	−4.99	−4.63	−3.62	−4.19	−1.65	−0.56	−0.16	0.36	0.85	1.35	1.88	2.36	2.86	3.29
	$\hat{S}_d$	0.01	0.25	0.28	1.17	0.28	0.30	0.29	0.29	0.28	0.28	0.29	0.28	0.29	0.28
0.2	$\hat{d}$	−4.40	−3.41	−2.41	−1.40	−0.41	0.57	1.08	1.58	2.10	2.58	3.08	3.60	4.11	4.51
	$\hat{S}_d$	0.28	0.27	0.26	0.26	0.27	0.28	0.26	0.26	0.27	0.28	0.27	0.26	0.26	0.26
0.4	$\hat{d}$	−2.85	−1.86	−0.85	−1.30	1.14	2.14	2.65	3.16	3.64	4.15	4.63	4.94	4.99	4.99
	$\hat{S}_d$	0.27	0.27	0.27	2.13	0.29	0.28	0.28	0.28	0.26	0.27	0.25	0.11	0.01	0.03
0.6	$\hat{d}$	−0.91	−0.09	1.08	0.73	3.08	4.10	4.59	4.92	4.99	4.99	4.99	4.99	4.99	4.99
	$\hat{S}_d$	0.28	0.29	0.30	2.23	0.29	0.29	0.27	0.14	0.02	0.01	0.00	0.00	0.00	0.00
0.8	$\hat{d}$	1.41	2.41	3.43	2.89	4.98	4.99	4.99	4.99	4.99	4.99	4.99	4.99	4.99	4.99
	$\hat{S}_d$	0.29	0.29	0.31	2.02	1.76	0.04	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.00

Note: The means and the standard deviations are the reported values multiplied by 10^{-1}.

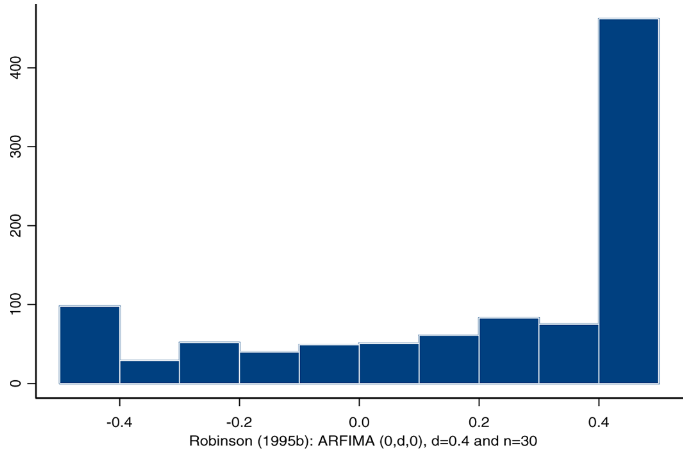

Figure 9. Robinson (1995b), $n = 30$

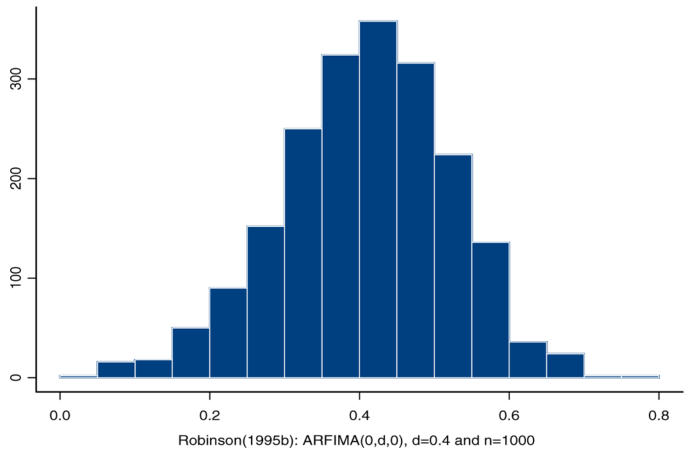

Figure 10. Robinson (1995b), $n = 1000$

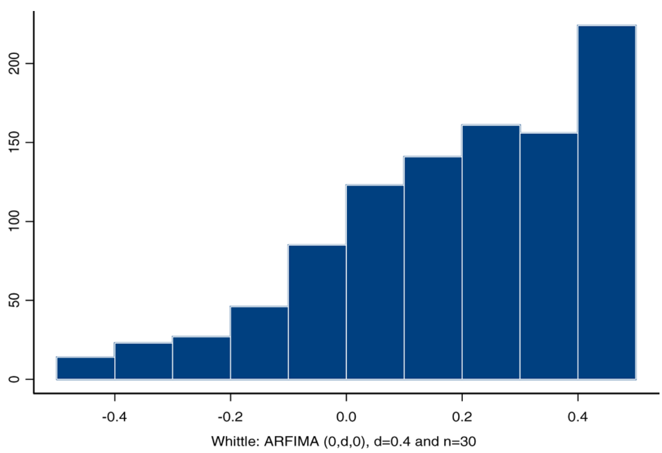

Figure 11. Whittle for $n = 30$

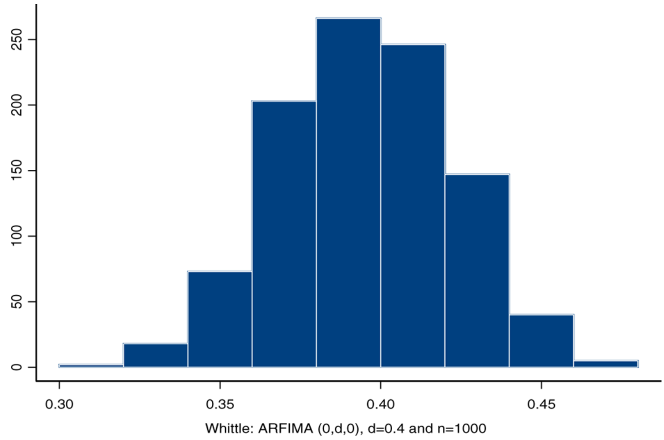

Figure 12. Whittle for $n = 1000$

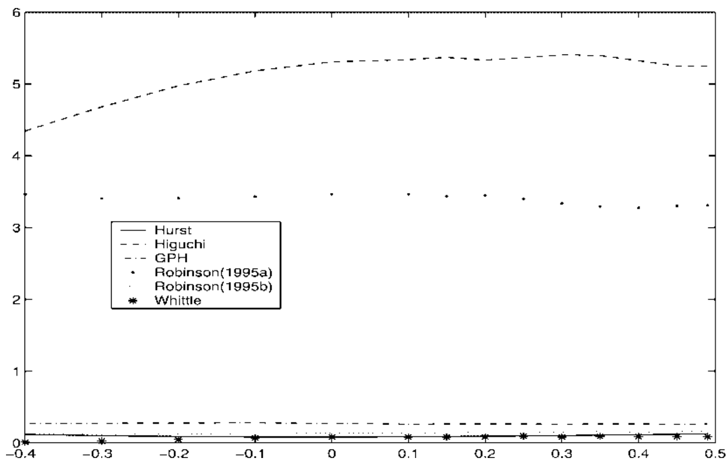

Figure 13. The MSE for ARFIMA $(0, d, 0)$ and $n = 30$

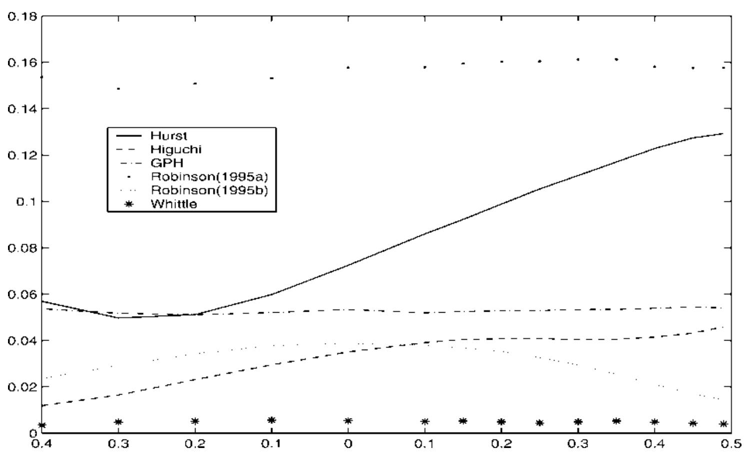

Figure 14. The MSE for ARFIMA $(0, d, 0)$ and $n = 200$

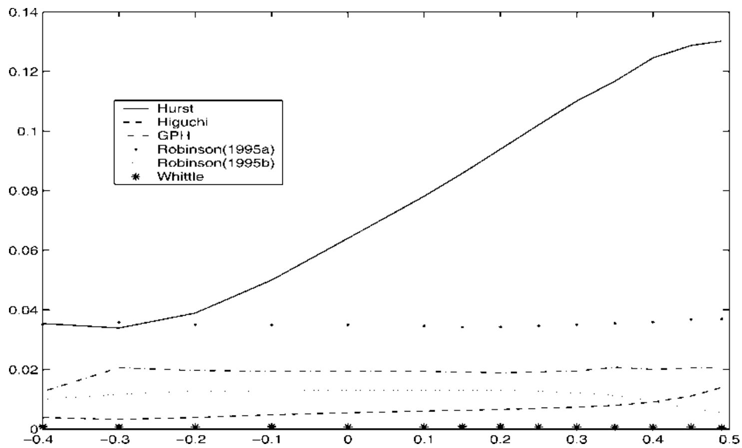

Figure 15. The MSE for ARFIMA (0, d, 0) and $n = 1000$

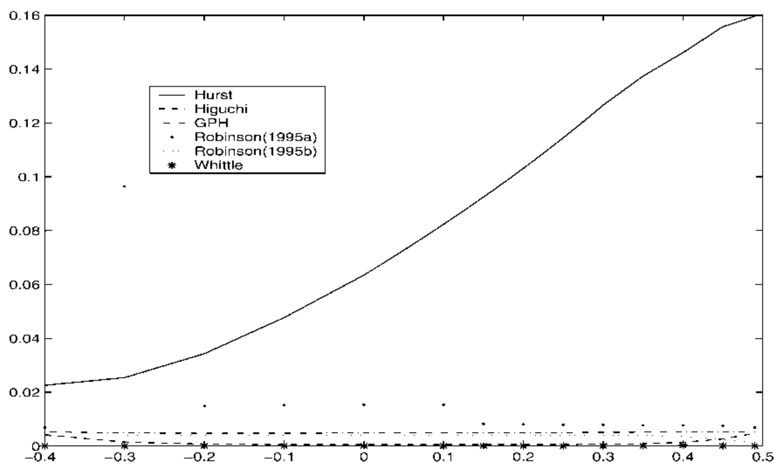

Figure 16. The MSE for ARFIMA (0, d, 0) and $n = 10,000$

Table 24. The results of the stationarity tests

Tests	ADF	PP	KPSS
π_t	−2.169	−9.381	0.857***

impact on the mean and the estimated standard deviation of the estimators obtained by the different methods. In this paper, our study is limited to a stationary case, i.e. $d < 0.5$. However, many economic time series are non-stationary. Some of methods considered in this paper are extended to the non-stationary case. Their performance must be examined in the same way in future work.

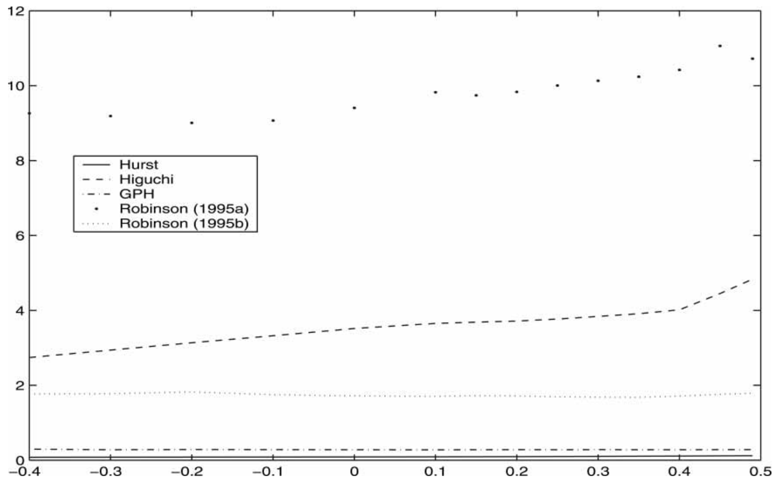

Figure 17. The MSE for FGN and $n = 30$

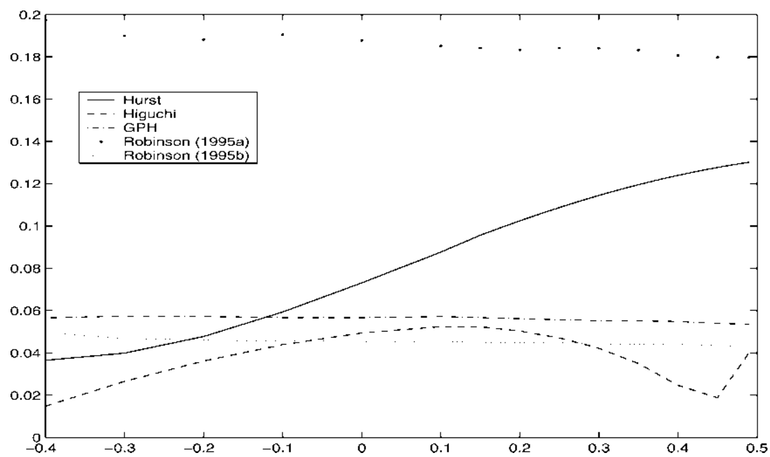

Figure 18. The MSE for FGN and $n = 200$

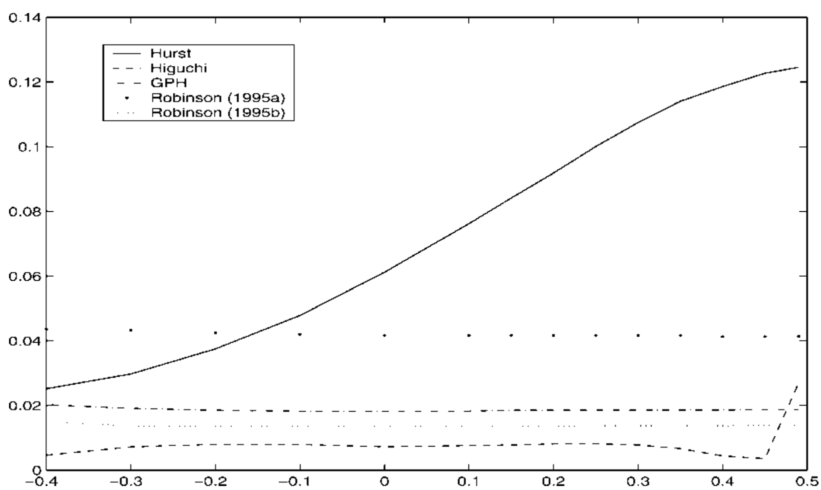

Figure 19. The MSE for FGN and $n = 1000$

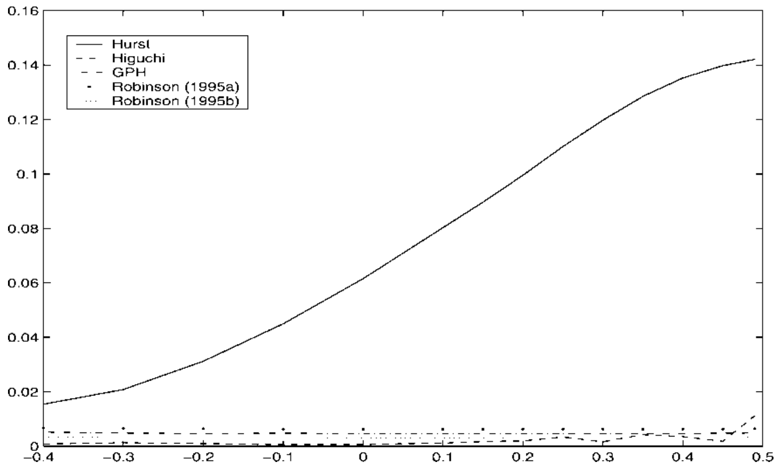

Figure 20. The MSE for FGN and $n = 10,000$

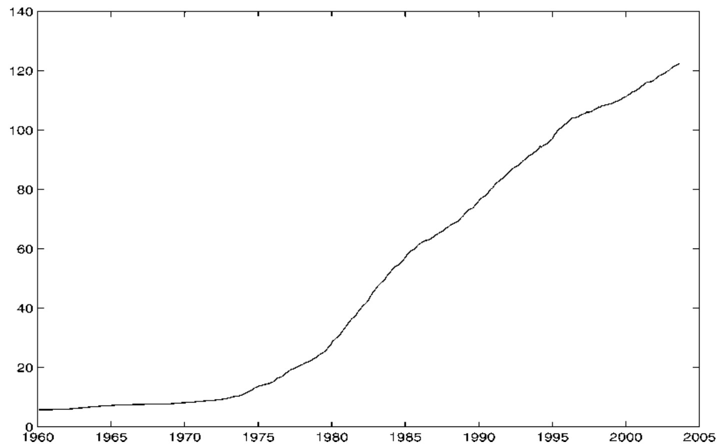

Figure 21. The LCPI series

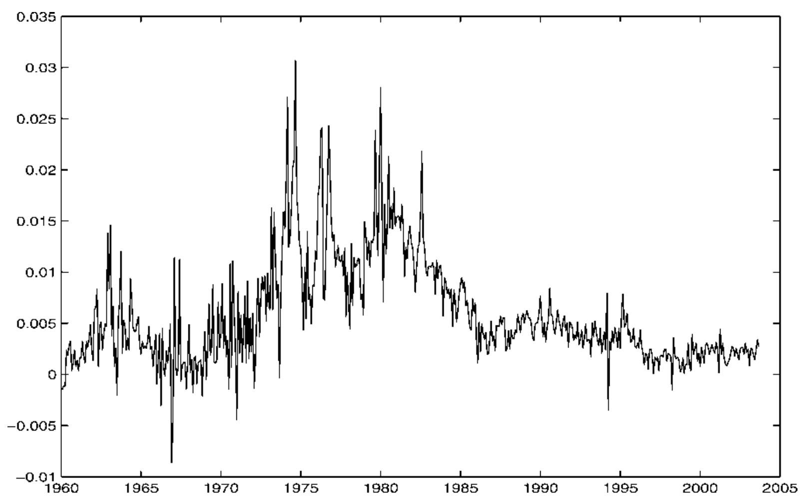

Figure 22. The SA US inflation rate

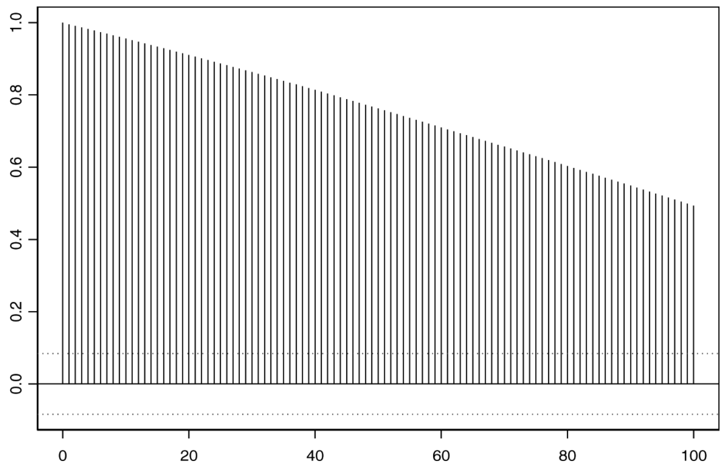

Figure 23. Correlogram of LCPI series

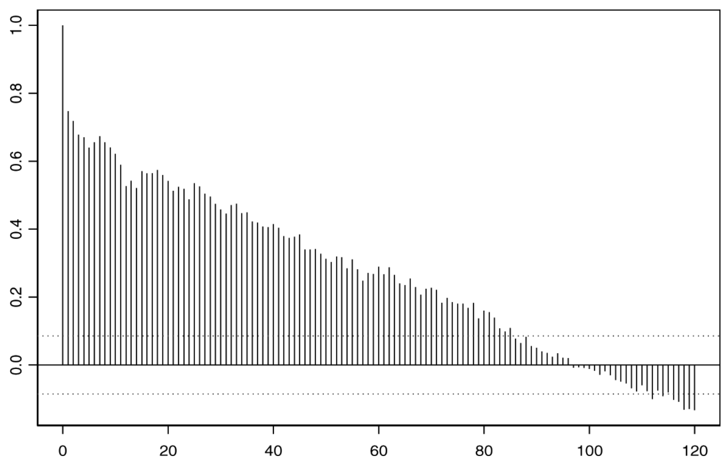

Figure 24. Correlogram of SA US Inflation rate

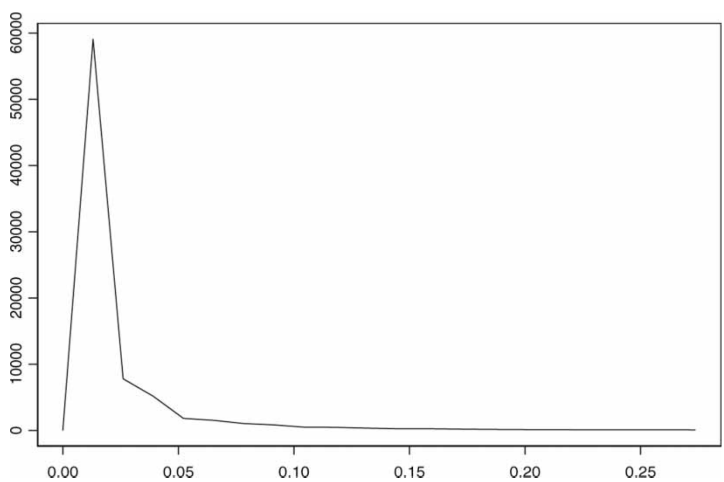

Figure 25. Periodogram of LCPI series

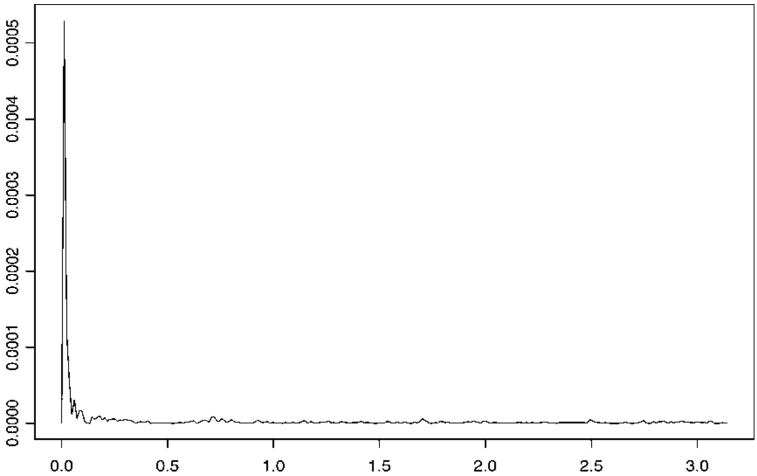

Figure 26. Periodogram of SA US inflation rate

Table 25. Estimates of long-range dependence parameter d

	Hurst	Higuchi	GPH	Robinson (1995a)	Robinson(1995b)	Whittle	MV
$\hat{d}$	0.532	0.993	0.456	0.371	0.489	0.445	0.448

Table 26. The results of specification tests of original residuals

	GPH	Robinson (1995a)	Robinson (1995b)	Whittle	MV
LB	39.496	56.996	41.417	39.682	39.582
p-values	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
SW	0.935	0.926	0.939	0.934	0.934
p-values	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Table 27. The results of specification tests of transformed residuals

	GPH	Robinson (1995a)	Robinson (1995b)	Whittle	MV
The order p	1	1	1	1	1
$\hat{\phi}_1$	0.301	0.294	0.300	0.304	0.301
LB	0.944	1.822	2.660	2.663	2.786
p-values	(1.000)	(0.999)	(0.997)	(0.997)	(0.997)
SW	0.944	0.942	0.945	0.944	0.944
p-values	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

Notes

[1]There are many classes of models that can characterize the long memory, such as the Gegenbauer processes, but in this paper we will consider only the fractional Gaussian noise and the ARFIMA (p, d, q) process.

[2]The parameter H is called the Hurst coefficient since it is originally due to Hurst (1951).

[3]It should be noted that the bias of the estimators obtained by the semi-parametric methods depends on the truncation parameter m. Geweke & Porter-Hudak (1983) suggested that $m = n^{1/2}$ is the best choice,

however, many other authors showed that this choice is not optimal, see for example, Hurvich & Beltrao (1993), Henry & Robinson (1996), Hurvich *et al.* (1998), Hurvich & Deo (1999).

[4] c is the Euler constant and equal to $0.57721\ldots$.

[5] Reisen (1994) proposed an estimator that is an extension of the GPH one. To determine this estimator, he used the smoothed periodogram. An application on simulated data showed that this estimator is better than the GPH one since it has a smaller bias and a smaller standard deviation.

[6] Fox & Taqqu (1986) showed theoretically the similarity of these two methods, therefore in the simulation study, we limit ourselves to the Whittle method.

[7] The application of the Hurst method and their extensions, the semi-parametric methods as well as the maximum likelihood methods with ARFIMA models is justified by Theorem 4.2 in Beran (1994) with $H = d + 1/2$.

[8] In order to examine the limit case of the stationarity, we run an additional experiment with $d = 0.49$.

[9] The simulations are performed with S-PLUS 6.0 and Gauss 3.0.

[10] When $\phi_1 \to 1$, $\{X_t\}$ becomes non-stationary and the results break down naturally.

[11] For the other methods, we use the same characteristics to plot the histograms, $n = 30$, $n = 1000$ and $d = 0.4$.

[12] The MSE are only given for an ARFIMA $(0, d, 0)$ process, in order not to take into account the short-memory effects of an ARFIMA (p, d, q) process.

[13] The histograms of the FGN process are available under request.

[14] The authors thank Bhardwaj, G. for the data basis.

[15] Note that, 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%). (∗∗∗) indicates that the hypothesis of stationarity is rejected at the 1% significant level.

References

Andrews, D. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation, *Econometrica*, 59, pp. 817–858.

Beran, J. (1994) *Statistics for Long-Memory Processes* (Chapman & Hall).

Beran, J. & Ocker, D. (1999) SEMIFAR forecasts, with applications to foreign exchanges rates, *Journal of Statistical Planning and Inference*, 80, pp. 137–153.

Bhardwaj, G. & Swanson, N. R. (2006) An empirical investigation of the usefulness of ARFIMA models for predicting macroeconomic and financial time series, *Journal of Econometrics*, 131, pp. 539–578.

Bos, C. S., Franses, P. H. & Ooms, M. (2002) Inflation, forecast intervals and long memory regression models, *International Journal of Forecasting*, 110, pp. 167–185.

Box, G. E. P. & Jenkins, G. M. (1976) *Time Series Analysis Forecasting and Control* (San Francisco: Holden Day).

Cheung, Y. W. (1993) Long memory in foreign-exchanges rates, *Journal of Business and Economic Statistics*, 11, pp. 93–101.

Cheung, Y.-W. & Lai, K. S. (1993) A fractional cointegration analysis of purchasing power parity, *Journal of Business and Economic Statistics*, 11, pp. 103–112.

Chio, K. & Zivot, E. (2002) Long memory and structural changes in the forward discount: an empirical investigation, *Working paper*, (University of Washington).

Chow, K. V., Denning, K. C., Ferris, S. & Noronha, G. (1995) Long-term and short-term price memory in the stock market, *Economics Letters*, 49, pp. 287–293.

Cox, D. R. (1989) Long-range dependence, non-linearity and time irreversibility, *Journal of Time Series Analysis*, 4, pp. 329–335.

Crato, N. & De Lima, P. J. F. (1994) Long-range dependence in the conditional variance of stocks returns, *Economics Letters*, 45, pp. 281–285.

Dahlhaus, R. (1989) Efficient parameter estimation for self-similar processes, *The Annals of Statistics*, 17, pp. 1749–1766.

Davies, R. B. & Harte, D. S. (1987) Tests for Hurst effect, *Biometrika*, 74, pp. 95–101.

Davidson, J. & Sibbertsen, P. (2005) Generating schemes for long memory processes: regimes, aggregation and linearity, *Journal of Econometrics*, 128, pp. 253–282.

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

Fox, R. & Taqqu, M. (1986) Large sample properties of parameter estimates for strongly dependent stationary time series, *The Annals of Statistics*, 14, pp. 517–532.

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

Granger, C. W. J. (1980) Long memory relationships and the aggregation of dynamic models, *Journal of Econometrics*, 14, pp. 227–238.

Granger, C. W. J. & Joyeux, R. (1980) An introduction to long-memory time series models and fractional differencing, *Journal of Time Series Analysis*, 1, pp. 15–29.

Grenander, U. & Szegö, G. (1958) *Toeplitz Forms and their Application* (University of California Press, Berkeley).

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

Henry, M. & Robinson, P. M. (1996) Bandwidth choice in Gaussian semiparametric estimation of long-range dependence (Proc). *Athens Conference on Applied Probability and Time Series*, edited by P. M. Robinson and M. Rosenblatt, pp. 220–232 (Taipei: Academia Sinica).

Higuchi, T. (1988) Approach to an irregular time series on the basis of the fractal theory, *Physica*, *D*, 31, pp. 277–283.

Hosking, J. R. M. (1981) Fractional Differencing, *Biometrika*, 68, pp. 165–176.

Hurst, H. E. (1951) Long-term storage capacity of reservoirs, *Transactions of the American Society of Civil Engineers*, 116, pp. 770–799.

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

Hurvich, C. M. & 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, *Journal of Time Series Analysis*, 3, pp. 331–341.

Hurvich, C. M., Deo, R. S. & 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, pp. 19–46.

Hyung, N. & Franses, P. H. (2001) Structural breaks and long memory in US inflation rates: do they matter for forecasting? *Econometric Institute Research Report*, 13.

Jeffreys, H. (1939) *Theory of Probability* (Oxford: Clarendon Press).

Jensen, M. J. (1994) *Wavelet Analysis of Fractionally Integrated Processes* (St. Louis, MO 63130: Department. of Economics, Washington University).

Kurpiel, A. & Marimoutou, V. (2000) Comparaison de tests de processus mémoire longue par Monte Carlo: Une approche en niveau et puissance. Unpublished paper.

Kwiatkowski, D., Phillips, P. C. B., 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, pp. 159–178.

Ljung, G. M. & Box, G. E. P. (1978) On a measure of lack of fit in time series models, *Biometrika*, 65, pp. 553–564.

Lo, A. (1991) Long-term memory in stock market prices, *Econometrica*, 59, pp. 1279–1313.

Lo, A. & MacKinlay, C. (1989) The size and power of the variance ratio test in finite samples: a Monte Carlo investigation, *Journal of Econometrics*, 40, pp. 203–238.

Lobato, I. & Robinson, P. M. (1996) Averaged periodogram estimation of long memory, *Journal of Econometrics*, 73, pp. 303–324.

Mandelbrot, B. B. (1962) Sur certains prix spéculatifs: faits empiriques et modèle basé sur les processus stables additifs de Paul Lévy, *Comptes Rendus* (Paris), 254, pp. 3968–3970.

Mandelbrot, B. B. (1971) A fast fractional Gaussian noise generator, *Water Resources Research*, 7, pp. 543–553.

Mandelbrot, B. B. & Van Ness, J. W. (1968) Fractional Brownian motions, fractional noises and applications, *SIAM Review*, 10, pp. 422–437.

Mandelbrot, B. B. & Taqqu, M. S. (1979) Robust R/S analysis of long-run serial correlation, *42nd Session of the ISI, Manilla*, pp. 4–14.

Newey, W. & West, K. (1987) A simple positive definite, heteroscedasticity and autocorrelation consistent covariance matrix, *Econometrica*, 55, pp. 703–705.

Newcomb, S. (1886) A generalized theory of the combination of observations so as to obtain the best result, *American Journal of Mathematics*, 8, pp. 343–366.

Phillips, P. C. B. (1987) Time series regression with a unit root, *Econometrica*, 55, pp. 277–301.

Phillips, P. C. B. & Perron, P. (1988) Testing for a unit root in time series regression, *Biometrika*, 75, pp. 335–346.

Porter-Hudak, S. (1990) An application to the seasonal fractionally differenced model to the monetary aggregates, *Journal of the American Statistical Association*, 85, pp. 338–344.

Reisen, V. A. (1994) Estimtaion of the fractional difference parameter in the ARIMA (p, d, q) model using the smoothed periodogram, *Journal of Time Series Analysis*, 15, pp. 335–350.

Robinson, P. M. (1978) Statistical inference for a random coefficient autoregressive model, *Scandinavian Journal of Statistics*, 5, pp. 163–168.

Robinson, P. M. (1994) Semiparametric analysis of long-memory time series, *The Annals of Statistics*, 22, pp. 515–539.

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

Robinson, P. M. (1995b) Gaussian semiparametric estimation of long range dependence, *The Annals of Statistics*, 23, pp. 1630–1661.

Sowell, F. (1992) Maximum likelihood estimation of stationary univariate fractionally integrated time series models, *Journal of Econometrics*, 53, pp. 165–188.

Stock, J. & Watson, M. (2002) Macroeconomic forecasting using diffusion indexes, *Journal of Business and Economic Statistics*, 20, pp. 147–162.

Student. (1927) Errors of routine analysis, *Biometrika*, 19, pp. 151–164.

Taqqu, M. S. & Teverovsky, V. (1998) On estimating the intensity of long-range dependence in finite and infinite variance time series, *A Practical Guide To Heavy Tails: Statistical Techniques and Applications*, pp. 177–217 (Boston).

Taqqu, M. S., Teverovsky, V. & Willinger, W. (1995) Estimators for long-range dependence: an empirical study, *Fractals*, 3, pp. 785–798.

Taqqu, M., Willinger, W. & Sherman, R. (1997) Proof of a fundamental result in self similar traffic modeling, *Computer Communication Review*, 27, pp. 5–23.

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

Whittle, P. (1951) *Hypothesis Testing in Time Series Analysis* (New York: Hafner).

Yajima, Y. (1985) On estimation of long-memory time series models, *Australian Journal of Statistics*, 27, pp. 303–320.