

Error and order misspecification in ARFIMA models

Valderio A. Reisen*

Manoel R. Sena Jr.**

Silvia R. C. Lopes***

Abstract

In developing the long and short memory estimation, it is usually assumed that the innovations in the ARFIMA model are normally distributed. However, circumstances may occur where this assumption is not true. This paper uses Monte Carlo simulation to evaluate the robustness of different estimators of the fractional parameter in stationary and invertible ARFIMA processes to the misspecification of the error distribution. In particular, we consider misspecification against heavy-tailed, skewed and bimodal distributions. The study is also extended for the incorrect ARFIMA specification.

Resumo

No processo de estimação dos parâmetros de longa e curta memória é usualmente assumido que os erros no modelo ARFIMA são normalmente distribuídos. Entretanto pode ocorrer, em certas circunstâncias, que esta suposição não seja verdadeira. Este artigo utiliza simulação de Monte Carlo para investigar a robustez de diferentes estimadores do parâmetro fracionário do processo ARFIMA, estacionário e invertível, quando os erros não são normais. Em particular, consideramos as distribuições de caldas pesadas, assimétricas e bimodais. O estudo é também estendido para o caso de ordem incorreta do processo ARFIMA.

Key Words: Fractional differencing, long memory, smoothed periodogram regression, periodogram regression, Whittle maximum likelihood procedure, non-normality, misspecification. .

JEL Code: 62M10, 62M15, 60G18 .

* Departamento de Estatística, CCE and PPGEA-CT-UFES, Vitória - ES, Brazil

** Departamento de Estatística, CCEN - UFPE, Recife - PE, Brazil

*** Departamento de Estatística, IM - UFRGS, Porto Alegre - RS, Brazil

1. Introduction.

Recently, inference issues with long memory time series have been considered in many works in the literature. The estimation of the long memory parameter has been carried out by using parametric procedures such as in Fox and Taqqu (1986), Dahlhaus (1989) and Sowell (1992) and semiparametric approaches such as in Geweke and Porter-Hudak (1983), Robinson (1994, 1995), Reisen (1994) and Lobato and Robinson (1996). Gaussianity has been assumed in a great deal of the work on long memory subjects. The parametric methods require the normality assumption for the errors of the innovations. However, this assumption may be relaxed in semiparametric approaches. This assumption is usually considered for the later methods, when one wants to analyze the limiting distribution properties of the estimators (see, for instance, Robinson (1995) and Lobato and Robinson (1996)).

The Gaussianity may not be a reasonable assumption in practical situations, specially in economic data. For example, as investigated by Anderson and Walker (1964), the stock price innovations often have highly leptokurtic distributions, see also Ljung and Box (1978). The later work also considers nonnormal errors in AR processes for testing the adequacy of the models. The interest in non-Gaussian errors for the ARFIMA processes has recently been considered by Giraitis and Taqqu (1999) and Velasco (2000). In the later work, the author shows the asymptotic properties for tapered and non tapered periodogram regression estimators. He assesses the validity of the asymptotic results for finite sample sizes, via Monte Carlo simulation, using the chi-square, exponential and t-Student distributions for the errors.

Our goal is to investigate, by Monte Carlo simulation, how the estimates of the fractional parameter d , based on parametric and semiparametric approaches, can be affected when the Gaussian error

assumption is violated.

In this sense, we want to verify how robust are these procedures (the word ‘robust’ is loaded with many -sometimes inconsistent- connotations). We shall follow the definition giving by Huber (1996) where ‘robustness’ signifies *insensitivity against small deviations from the assumptions*.

The following non-Gaussian innovation distributions are considered: exponential with parameter $\lambda = 1$ (recentered with zero mean), t-Student with three degrees of freedom, \mathcal{X}_1^2 (recentered with zero mean) and Gaussian mixture. The effect of small and large sample sizes is also investigated. The paper is organized as follows: in Section 2 the ARFIMA and the estimates of d are briefly presented. Section 3 gives the Monte Carlo simulation results and some concluding remarks are presented in Section 4.

2. The ARFIMA (p,d,q) Model.

The zero mean discrete time ARFIMA (p, d, q) model is given by

$$\Phi(\mathcal{B})(1 - \mathcal{B})^d X_t = \Theta(\mathcal{B}) \epsilon_t, \quad d \in \mathbb{R}, \text{ for all } t \in \mathbb{Z} \quad (2.1)$$

where \mathcal{B} is the lag operator, $\Phi(\mathcal{B}) = 1 - \phi_1 \mathcal{B} - \dots - \phi_p \mathcal{B}^p$ and $\Theta(\mathcal{B}) = 1 - \theta_1 \mathcal{B} - \dots - \theta_q \mathcal{B}^q$ are polynomials of orders p and q , respectively, with roots outside of the unit circle and $\epsilon_t \sim WN(0, \sigma_\epsilon^2)$.

When $d \in (-0.5, 0.5)$, the ARFIMA process is stationary and invertible, and its spectral density function, $f_X(\cdot)$, is given by

$$f_X(w) = f_U(w) \left[2 \sin\left(\frac{w}{2}\right) \right]^{-2d}, \quad w \in [-\pi, \pi], \quad (2.2)$$

where the function $f_U(\cdot)$ is the spectral function of an ARMA(p, q) process. Hosking (1981), Beran (1994) and Reisen (1994) describe ARFIMA models in detail.

We consider five alternative estimators for the parameter d . These estimators are briefly described below.

The first estimator, denoted hereafter by GPH, was proposed by Geweke and Porter-Hudak (1983) who uses the periodogram function $I(\cdot)$ as an estimate of the spectral density function in expression (2.2). The number of observations in the regression equation is a function of the sample size n , that is, $g(n) = n^\alpha$, $0 < \alpha < 1$.

The second estimator denoted in the sequel by SPER was introduced by Reisen (1994). This regression estimator is obtained by replacing the spectral density function in the expression (2.2) by the smoothed periodogram function based on the Parzen lag window. In this estimation method $g(n)$ is chosen as above and the truncation point in the Parzen lag window is $m = n^\beta$, $0 < \beta < 1$. The appropriate choices of α and β were investigated by Geweke and Porter-Hudak (1983) and Reisen (1994), respectively.

The third estimator, suggested by Robinson (1995) and denoted hereafter by ROB, is a modified version of the Geweke and Porter-Hudak's estimator. This estimator regresses $\ln\{I(w_i)\}$ on $\ln\{2\sin(w_i/2)\}^2$, for $i = l, l+1, \dots, g(n)$, where l is the lower truncation point. In the asymptotic theory both l and $g(n)$ tend to infinity with n , but more slowly, while $\frac{l}{g(n)}$ goes to zero. Robinson (1995) derives some asymptotic results for his estimator when $d \in (-0.5, 0.5)$. Our choice of the bandwidth $g(n)$ is also given as above.

The appropriate choice of the optimal $g(n)$ has been the subject of many papers such as Hurvich et al. (1998) and Hurvich and Deo (1999).

The fourth semiparametric estimator is the one suggested by Robinson (1994) and Lobato and Robinson (1996), denoted hereafter by LOB. It is the weighted averages of the unlogged periodogram.

This estimator is based on the number of frequencies τ and on a constant $q \in (0.0, 1.0)$. Lobato and Robinson (1996) presented a Monte Carlo simulation study to investigate the sensitivity on the choice of τ and q .

The fifth estimator is a parametric procedure suggested by Fox and Taqqu (1986) based on Whittle (1953) and will be denoted hereafter by FT. The estimator FT is based on the periodogram and the spectral density functions. For computational purposes the estimator FT is obtained by minimizing the finite and discrete form given by

$$\mathcal{L}_n(\zeta) = \frac{1}{2n} \sum_{j=1}^{n-1} \left\{ \ln f_X(w_j, \zeta) + \frac{I(w_j)}{f_X(w_j, \zeta)} \right\}, \quad (2.3)$$

where ζ denotes the vector of unknown parameters (see Dahlhaus (1989), page 1753).

An extensive simulation study comparing parametric and semi-parametric methods to estimate the parameters of stationary Gaussian ARFIMA process is recently given by Smith et al. (1997), Bisaglia and Guégan (1998), Reisen and Lopes (1999) and Reisen et al. (2000), Reisen et al (2001) among others.

In the context of parametric methods we did not use the exact maximum likelihood. Moreover, evidence suggests that the performances of Sowell's exact maximum likelihood and Whittle's approximated procedure yield very similar results in the realistic case of unknown estimated mean for the ARFIMA(0, d , 0) model (see Cheung and Diebold (1994)).

3. Monte Carlo Simulation Study.

This estimation simulation study is divided into three situations: (1) when the generated process is the pure ARFIMA(0, d , 0); (2) when short-run coefficients are introduced in the model and (3) when the model is misspecified. These investigations are presented in the subsections 3.1, 3.2 and 3.3, respectively. All simulated ARFIMA times series throughout the paper were generated via the algorithm suggested by Hosking (1984) for sample size $n = 100, 300$ and the estimates are the average of 1,000 replications. The random numbers were obtained from the subroutine RNNOR in the IMSL library. The non-Gaussian innovations have the form: Gaussian Mixture $\mathcal{N}(0.707, 0.5)$ and $\mathcal{N}(-0.707, 0.5)$ with mixing proportion given by 0.5 (see Little and Rubin (1987)), Exponential with parameter $\lambda = 1$ (recentered with zero mean), t-Student with three degrees of freedom and Chi-square with one degree of freedom (recentered with zero mean). The Gaussian distribution is also considered for comparison purposes. For the semiparametric estimators we consider two different values for the bandwidth $g(n)$, that is, $g(n) = n^{0.5}, n^{0.7}$. For the smoothed periodogram estimator, SPER, we use $\beta = 0.9$ in the Parzen lag window as suggested by Reisen (1994). The lower truncation point l in the ROB method was set equal to 2 and in the LOB method we consider $q = 0.5$ and $\tau = n^{0.7}, n^{0.8}$ (see Lobato and Robinson (1996)). In the Whittle's method, the value of the estimates are obtained by using the subroutine BCONF in the same library. One will find in the tables: the mean estimated value of the parameter d ($mean(\hat{d})$), its bias ($bias$), its standard deviation (sd) and its mean squared error (mse).

We have considered several values of d in the range (-0.5,0.5),

i.e., antipersistent processes ($d < 0.0$) and persistent or long memory processes ($d > 0.0$). However, we report here only the results for $d = 0.2, 0.4$ since the pattern is similar for the other cases and they are available upon request.

3.1 The ARFIMA $(0, d, 0)$.

The results are presented in Tables 1-3. The effect of the sample sizes is evident, increasing n the estimates are much better for any innovation distribution considered (compare Tables 1 and 2). The same pattern was observed when $n = 100$ and $d = 0.4$ and this result is not presented here but is available from the authors. In spite of the error distribution, the Whittle's method (FT) is always superior in terms of estimating the parameter d . It is more accurate than the others methods, since we are dealing with the correct process order. However, the semiparametric methods also give good results. Comparing the log-regression estimators, GPH, SPER and ROB, the later has, in general, larger sd and mse . SPER presents smaller mse . The methods GPH and ROB, in general, overestimate the true parameter value while it is underestimated for SPER. Increasing the bandwidth $g(n)$ the estimates improved substantially by showing smaller $bias$ and mse . This was expected since these methods are obtained from an approximated regression equation and, in the case of the ARFIMA $(0, d, 0)$ model, the plot of $\ln\{\hat{f}\}$ against $\ln\{2\sin(w_i/2)\}^2$ is practically linear where $\{\hat{f}\}$ is the estimate of the spectral density function used in each method. The method LOB tends to underestimate the parameter with large bias compared with the others semiparametric methods. Increasing τ the LOB estimates do not improve substantially, and the bias also increases with d (see

Table 3). Apart from the estimation behaviour of LOB, this method has also the difficulty of choosing the appropriate values of q and τ . We have also considered $\tau = n^{0.5}$. However, the results are worse than those presented in the tables.

All estimators are not very sensitive to the type of the errors distribution considered since the estimation results from the non-Gaussian errors are very close to those obtained under the normality assumption.

Figures 1 and 2 present a graphical picture of the Monte Carlo and Normal distributions of the estimates when the innovations are respectively from the Normal and Normal mixture distributions. The others cases shown the same pattern and are not presented here to save space but are available upon request. The graphs confirm what we reported from the results presented in Table 2. It is not visible the difference among the estimates when considering Gaussian and non-Gaussian error distributions. Hence in the case of the ARFIMA $(0, d, 0)$ the parametric and semiparametric approaches are not affected when the Gaussian error assumption is violated.

Table 1: Simulation for the ARFIMA(0,0.2,0) model with different error distributions and $n = 100$

Error Distribution: Normal									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1983	0.1963	0.1269	0.1609	0.2013	0.1964	0.1196	0.1392	0.1937
$bias$	-0.0017	-0.0037	-0.0731	-0.0391	0.0013	-0.0036	-0.0804	-0.0608	-0.0063
sd	0.2950	0.1578	0.2182	0.1282	0.4169	0.1970	0.1228	0.0933	0.0945
mse	0.0869	0.0249	0.0529	0.0179	0.1736	0.0388	0.0215	0.0124	0.0090
Error Distribution: Normal Mixture									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2166	0.2156	0.1270	0.1759	0.2275	0.2180	0.1374	0.1412	0.1983
$bias$	0.0166	0.0156	-0.0730	-0.0241	0.0275	0.0180	-0.0626	-0.0588	-0.0017
sd	0.2791	0.1541	0.2051	0.1230	0.4025	0.1933	0.1195	0.1008	0.0930
mse	0.0781	0.0240	0.0473	0.0157	0.1626	0.0377	0.0182	0.0136	0.0086
Error Distribution: Exponential									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2179	0.2086	0.1385	0.1672	0.2095	0.2038	0.1258	0.1420	0.2000
$bias$	0.0179	0.0086	-0.0615	-0.0328	0.0095	0.0038	-0.0742	-0.0580	0.0000
sd	0.2929	0.1580	0.2219	0.1290	0.4292	0.1996	0.1163	0.0898	0.0922
mse	0.0860	0.0250	0.0530	0.0177	0.1841	0.0398	0.0190	0.0114	0.0085
Error Distribution: t-Student									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1944	0.1933	0.1265	0.1619	0.1931	0.1924	0.1217	0.1384	0.1778
$bias$	-0.0056	-0.0067	-0.0735	-0.0381	-0.0069	-0.0076	-0.0783	-0.0616	-0.0222
sd	0.2884	0.1610	0.2121	0.1294	0.4330	0.2047	0.1197	0.0978	0.0978
mse	0.0831	0.0259	0.0503	0.0182	0.1873	0.0419	0.0204	0.0134	0.0100
Error Distribution: Qui-Square									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2133	0.2120	0.1264	0.1720	0.2073	0.2094	0.1345	0.1402	0.1820
$bias$	0.0133	0.0120	-0.0736	-0.0280	0.0073	0.0094	-0.0655	-0.0598	-0.0180
sd	0.2751	0.1565	0.2048	0.1220	0.4136	0.1963	0.1070	0.0896	0.0902
mse	0.0758	0.0246	0.0473	0.0156	0.1709	0.0386	0.0157	0.0116	0.0085

Table 2: Simulation for the ARFIMA(0,0.2,0) model with different error distributions and $n = 300$

Error Distribution: Normal									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2060	0.1989	0.1551	0.1830	0.2153	0.2006	0.1699	0.1668	0.1981
$bias$	0.0060	-0.0011	-0.0449	-0.0170	0.0153	0.0006	-0.0301	-0.0332	-0.0019
sd	0.2118	0.1004	0.1604	0.0811	0.2668	0.1153	0.0744	0.0583	0.0483
mse	0.0448	0.0101	0.0277	0.0069	0.0713	0.0133	0.0064	0.0045	0.0023
Error Distribution: Normal Mixture									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2069	0.2008	0.1498	0.1842	0.2079	0.2000	0.1713	0.1661	0.1954
$bias$	0.0069	0.0008	-0.0502	-0.0158	0.0079	0.0000	-0.0287	-0.0339	-0.0046
sd	0.2043	0.0987	0.1598	0.0821	0.2709	0.1139	0.0729	0.0573	0.0473
mse	0.0418	0.0097	0.0280	0.0070	0.0734	0.0130	0.0061	0.0044	0.0023
Error Distribution: Exponential									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1891	0.1989	0.1406	0.1795	0.2011	0.2038	0.1685	0.1626	0.1957
$bias$	-0.0109	-0.0011	-0.0594	-0.0205	0.0011	0.0038	-0.0315	-0.0374	-0.0043
sd	0.2073	0.1008	0.1610	0.0841	0.2627	0.1150	0.0760	0.0570	0.0511
mse	0.0430	0.0102	0.0294	0.0075	0.0689	0.0132	0.0068	0.0046	0.0026
Error Distribution: t-Student									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1850	0.1973	0.1386	0.1784	0.1863	0.1994	0.1681	0.1665	0.1920
$bias$	-0.0150	-0.0027	-0.0614	-0.0216	-0.0137	-0.0006	-0.0319	-0.0335	-0.0080
sd	0.1987	0.1050	0.1575	0.0856	0.2658	0.1198	0.0756	0.0565	0.0491
mse	0.0397	0.0110	0.0285	0.0078	0.0708	0.0143	0.0067	0.0043	0.0025
Error Distribution: Qui-Square									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2029	0.2035	0.1563	0.1840	0.2061	0.2044	0.1717	0.1675	0.1927
$bias$	0.0029	0.0035	-0.0437	0.0160	0.0061	0.0044	-0.0283	-0.0325	-0.0073
sd	0.1951	0.0970	0.1514	0.0782	0.2472	0.1120	0.0757	0.0546	0.0537
mse	0.0380	0.0094	0.0248	0.0064	0.0611	0.0125	0.0065	0.0040	0.0029

Table 3: Simulation for the ARFIMA(0,0.4,0) model with different error distributions and $n = 300$

Error Distribution: Normal									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.4081	0.4080	0.3539	0.3879	0.4089	0.4081	0.3210	0.3228	0.4020
$bias$	0.0081	0.0080	-0.0461	-0.0121	0.0089	0.0081	-0.0790	-0.0772	0.0020
sd	0.2036	0.1002	0.1626	0.0830	0.2611	0.1121	0.0522	0.0446	0.1104
mse	0.0415	0.0101	0.0285	0.0070	0.0682	0.0126	0.0090	0.0079	0.0122
Error Distribution: Normal Mixture									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.4148	0.4119	0.3566	0.3913	0.4181	0.4122	0.3223	0.3231	0.4062
$bias$	0.0148	0.0119	-0.0434	-0.0087	0.0181	0.0122	-0.0777	-0.0769	0.0062
sd	0.2113	0.0987	0.1647	0.0825	0.2641	0.1125	0.0515	0.0425	0.1077
mse	0.0448	0.0099	0.0290	0.0069	0.0700	0.0128	0.0087	0.0077	0.0116
Error Distribution: Exponential									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.4072	0.4077	0.3528	0.3884	0.3962	0.4046	0.3195	0.3231	0.4090
$bias$	0.0072	0.0077	-0.0472	-0.0116	-0.0038	0.0046	-0.0805	-0.0769	0.0090
sd	0.2051	0.1023	0.1657	0.0841	0.2642	0.1147	0.0541	0.0430	0.1061
mse	0.0421	0.0105	0.0297	0.0072	0.0698	0.0132	0.0094	0.0078	0.0113
Error Distribution: t-Student									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.4174	0.4126	0.3584	0.3910	0.4191	0.4124	0.3207	0.3252	0.4024
$bias$	0.0174	0.0126	-0.0416	-0.0090	-0.0191	-0.0124	-0.0793	-0.0748	0.0024
sd	0.2044	0.1013	0.1615	0.0824	0.2663	0.1147	0.0538	0.0422	0.1092
mse	0.0420	0.0104	0.0278	0.0069	0.0713	0.0133	0.0092	0.0074	0.0119
Error Distribution: Qui-Square									
Estimate	GPH($g(n)$)		SPER($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.4087	0.4025	0.3526	0.3864	0.3999	0.3990	0.3182	0.3236	0.4095
$bias$	0.0087	0.0025	-0.0474	-0.0136	-0.0001	-0.0010	-0.0818	-0.0764	0.0095
sd	0.1946	0.0986	0.1573	0.0822	0.2580	0.1136	0.0529	0.0409	0.1074
mse	0.0379	0.0097	0.0270	0.0069	0.0665	0.0129	0.0095	0.0075	0.0116

Figure 1: Monte Carlo and Normal distributions of the estimates of d with Normal innovations; ARFIMA(0,0.2,0), $n = 300$.

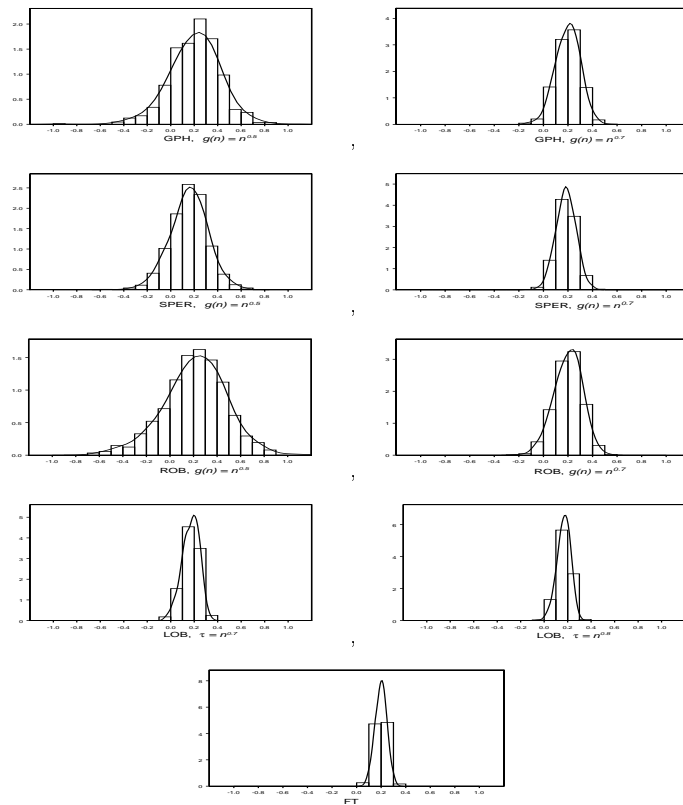
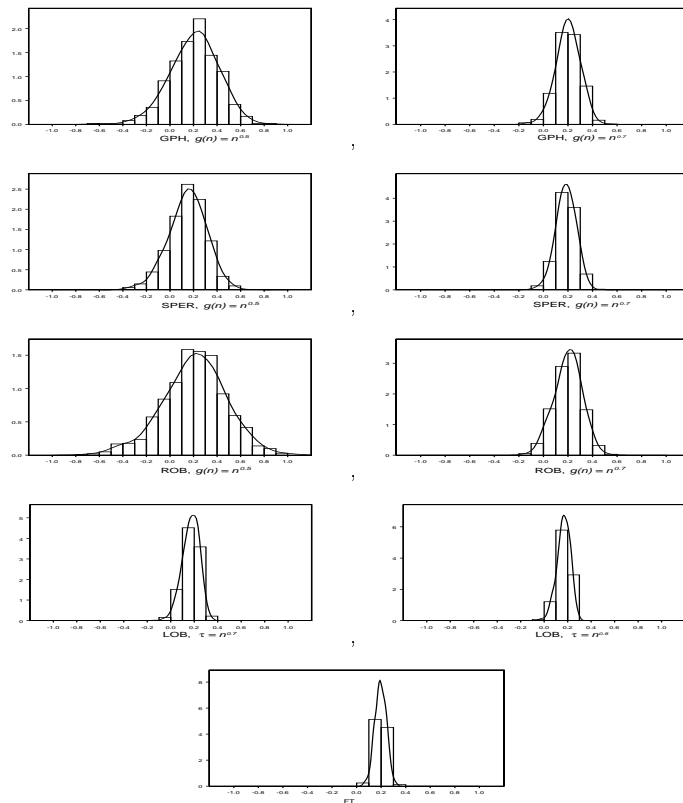


Figure 2: Monte Carlo and Normal distributions of the estimates of d with Normal Mixture innovations; ARFIMA(0,0.2,0), $n = 300$.



3.2 The ARFIMA (p, d, q) .

In this section we concentrate on the bias and on the mse of the fractional and short-run parameters of the ARFIMA (p, d, q) model, and the results are in Tables 4-11. This investigation shows the effect of the estimates when the short-run parameters are in the model. For the semiparametric methods the estimates of the short-run parameters are obtained in a staged process: first estimates the fractional parameter, then filters the series using this parameter estimate and finally estimates the appropriate AR and/or MA parameters by using the subroutine NSLE in the IMSL-Library. In this simulation, we assume that the true order is known and only the parameters need to be estimated. We have considered $\phi, \theta \in \{-0.2, -0.6, 0.2, 0.6\}$ and to save space we present the results when $n = 300$, $d = 0.2$ and $\phi, \theta \in \{-0.6, 0.6\}$. For $n = 100$ and $d = 0.4$ the pattern is similar and they are available upon request. The tables present the mean squared error of the long-run and short-run parameters estimates.

The ARFIMA $(1, d, 0)$

For $\phi = -0.6$ (see Table 4) the FT method is still superior. However, the methods GPH, ROB and SPER also perform well. The later approach outperforms the others two log-regression estimators. Increasing $g(n)$ the estimates get worse by showing larger bias and mean squared error. This occurs due to the fact of the strengthening contamination from the negative short-run parameter in the spectral density function. LOB is substantially inferior compared with all methods. This estimator presents very large bias specially when $\tau = n^{0.8}$. There is no evidence of the effect of non-Gaussian errors in the estimates.

The picture of the estimates changes when we are dealing with positive AR coefficients. The results are in Table 5. All methods are very sensitive for positive and large ϕ . Now FT is outperformed by the others methods. The best estimates are from SPER method

when $g(n) = n^{0.5}$ since we are using the smoothed periodogram as the estimate of the spectral density function. Again, increasing $g(n)$ or τ the bias of the semiparametric estimators increase substantially in contrast with ARFIMA(0, d , 0). This may be explained by the fact that the AR component enlarges the value of the spectral density function for frequencies away from the zero frequency.

In general for negative or positive values of the AR parameter, the bias in \hat{d} is compensated by the bias of $\hat{\phi}$. See, for instance, results in Table 5 ($\phi > 0$). While d is overestimated, the parameter ϕ is underestimated.

For positive ϕ , the presence of non-Gaussian error makes some changes in the estimates obtained from FT approach. For this method the difference between Gaussian and non-Gaussian estimates is large specially when considering the Qui-squared error distribution.

The ARFIMA (0, d , 1)

Tables 6-7 report the results for the ARFIMA (0,0.2,1). Now, large impact in the FT estimates occurs for negative values of θ in contrast with the negative values of ϕ in the ARFIMA (1,0.2,0) model, being this approach outperformed by the semiparametric methods (see Table 6). By comparing ARFIMA (0,0.2,1) and ARFIMA (1,0.2,0) the semiparametric estimators present much larger bias for $\theta > 0$ than for $\phi > 0$. This can be easily seen by the shape of the spectral density function in the neighborhood of zero frequency when $\theta > 0$. The estimates from LOB method become increasingly negative.

There is impact on the FT estimates for t-Student and Qui-square errors distributions when $\theta = 0.6$ as we also observed in ARFIMA(1, d , 0) case.

Table 4: Simulation for the ARFIMA(1,0.2,0) model with different error distributions, where $\phi = -0.6$ and $n = 300$.

Error Distribution: Normal									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1993	0.1558	0.1479	0.1378	0.1927	0.1460	0.1219	-0.0091	0.1936
mse	0.0365	0.0111	0.0265	0.0103	0.0662	0.0150	-0.0128	0.0500	0.0038
$mean(\hat{\phi})$	-0.5715	-0.5685	-0.5529	-0.5607	-0.5390	-0.5595	-0.5500	-0.4501	-0.5953
mse	0.0168	0.0061	0.0138	0.0061	0.0482	0.0084	0.0078	0.0300	0.0059
Error Distribution: Normal Mixture									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1943	0.1537	0.1457	0.1369	0.1891	0.1451	0.1244	-0.0076	0.1960
mse	0.0400	0.0126	-0.0277	0.0108	0.0692	0.0161	-0.0131	0.0498	0.0041
$mean(\hat{\phi})$	-0.5681	-0.5695	-0.5527	-0.5627	-0.5332	-0.5608	-0.5542	-0.4532	-0.6022
mse	0.0206	0.0065	0.0151	0.0059	0.0538	0.0092	0.0073	0.0290	0.0072
Error Distribution: Exponential									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1925	0.1603	0.1472	0.1432	0.1939	0.1546	0.1298	-0.0051	0.1996
mse	0.0416	0.0123	0.0266	0.0099	0.0782	0.0169	0.0116	0.0479	0.0039
$mean(\hat{\phi})$	-0.5639	-0.5727	-0.5541	-0.5665	-0.5287	-0.5665	-0.5578	-0.4562	-0.6038
mse	0.0247	0.0067	0.0152	0.0058	0.0728	0.0092	0.0068	0.0274	0.0074
Error Distribution: t-Student									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1975	0.1595	0.1413	0.1404	0.1971	0.1525	0.1266	-0.0045	0.2026
mse	0.0426	0.0120	0.0297	0.0104	0.0726	0.0150	0.0122	0.0474	0.0058
$mean(\hat{\phi})$	-0.5665	-0.5715	-0.5470	-0.5638	-0.5411	-0.5649	-0.5547	-0.4555	-0.6369
mse	0.0226	0.0069	0.0172	0.0062	0.0462	0.0089	0.0074	0.0281	0.0250
Error Distribution: Qui-Square									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1951	0.1584	0.1480	0.1408	0.1983	0.1525	0.1271	-0.0053	0.1995
mse	0.0408	0.0118	0.0271	0.0097	0.0719	0.0152	0.0118	0.0481	0.0054
$mean(\hat{\phi})$	-0.5640	-0.5700	-0.5520	-0.5632	-0.5332	-0.5638	-0.5541	-0.4533	-0.6300
mse	0.0231	0.0066	0.0153	0.0059	0.0639	0.0083	0.0073	0.0284	0.0231

Table 5: Simulation for the ARFIMA(1,0.2,0) model with different error distributions, where $\phi = 0.6$ and $n = 300$.

Error Distribution: Normal									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2652	0.4683	0.2118	0.4504	0.2840	0.5090	0.3584	0.4153	0.4957
mse	0.0459	0.0823	0.0235	0.0693	0.0740	0.1086	0.0265	0.0467	0.2211
$mean(\hat{\phi})$	0.5162	0.3335	0.5675	0.3496	0.4991	0.2966	0.4388	0.3862	0.3272
mse	0.0403	0.0815	0.0211	0.0700	0.0290	0.0479	0.0616	0.1044	0.1759
Error Distribution: Normal Mixture									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2621	0.4687	0.2184	0.4518	0.2766	0.5089	0.3585	0.4147	0.5497
mse	0.0457	0.0826	0.0243	0.0703	0.0729	0.1091	0.0268	0.0465	0.2542
$mean(\hat{\phi})$	0.5164	0.3295	0.5593	0.3446	0.5008	0.2929	0.4356	0.3832	0.2800
mse	0.0407	0.0835	0.0222	0.0723	0.0610	0.1069	0.0299	0.0491	0.2039
Error Distribution: Exponential									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2685	0.4664	0.2177	0.4497	0.2818	0.5045	0.3572	0.4140	0.5270
mse	0.0440	0.0803	0.0251	0.0688	0.0747	0.1057	0.0263	0.0462	0.2331
$mean(\hat{\phi})$	0.5125	0.3328	0.5603	0.3478	0.4992	0.2983	0.4375	0.3845	0.2995
mse	0.0391	0.0807	0.0223	0.0705	0.0608	0.1028	0.0292	0.0485	0.1874
Error Distribution: t-Student									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2621	0.4713	0.2096	0.4509	0.2770	0.5117	0.3599	0.4145	0.5331
mse	0.0450	0.0834	0.0261	0.0696	0.0720	0.1100	0.2676	0.0271	0.0464
$mean(\hat{\phi})$	0.5179	0.3278	0.5665	0.3459	0.5034	0.2909	0.4347	0.3837	0.2963
mse	0.0401	0.0840	0.0229	0.0718	0.0600	0.1074	0.0301	0.0489	0.2100
Error Distribution: Qui-Square									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2706	0.4769	0.2188	0.4549	0.2836	0.5164	0.3606	0.4162	0.5636
mse	0.0437	0.0854	0.0242	0.0708	0.0762	0.1120	0.0271	0.0471	0.2837
$mean(\hat{\phi})$	0.5161	0.3268	0.5643	0.3464	0.5008	0.2908	0.4379	0.3860	0.2725
mse	0.0387	0.0840	0.0209	0.0713	0.0594	0.1071	0.0291	0.0479	0.2218

Table 6: Simulation for the ARFIMA(0,0.2,1) model with different error distributions, where $\theta = -0.6$ and $n = 300$.

Error Distribution: Normal									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2114	0.2450	0.1620	0.2256	0.2081	0.2505	0.2092	0.2829	0.4109
mse	0.0388	0.0131	0.0241	0.0073	0.0658	0.0165	0.0052	0.0086	0.1545
$mean(\hat{\theta})$	-0.5861	-0.6145	-0.5738	-0.5849	-0.5798	-0.5701	-0.5943	-0.5581	-0.4268
mse	0.0135	0.0079	0.0067	0.0047	0.0254	0.0082	0.0037	0.0051	0.1044
Error Distribution: Normal Mixture									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2071	0.2428	0.1580	0.2227	0.2151	0.2519	0.2091	0.2795	0.3895
mse	0.0433	0.0109	0.0258	0.0065	0.0678	0.0140	0.0046	0.0081	0.1396
$mean(\hat{\theta})$	-0.5851	-0.5749	-0.6146	-0.5855	-0.5749	-0.5698	-0.5930	-0.5585	-0.4462
mse	0.0154	0.0052	0.0080	0.0039	0.0256	0.0063	0.0033	0.0046	0.0895
Error Distribution: Exponential									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2107	0.2430	0.1618	0.2259	0.2101	0.2486	0.2093	0.2800	0.3873
mse	0.0401	0.0120	0.0244	0.0070	0.0696	0.0153	0.0047	0.0083	0.1362
$mean(\hat{\theta})$	-0.5850	-0.5737	-0.6136	-0.5840	-0.5790	-0.5702	-0.5935	-0.5586	-0.4503
mse	0.0152	0.0067	0.0082	0.0045	0.0261	0.0078	0.0035	0.0048	0.0877
Error Distribution: t-Student									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2069	0.2478	0.1552	0.2285	0.2011	0.2538	0.2108	0.2816	0.4247
mse	0.0395	0.0120	0.0258	0.0072	0.0673	0.0151	0.0047	0.0085	0.1697
$mean(\hat{\theta})$	-0.5886	-0.5728	-0.6178	-0.5834	-0.5852	-0.5686	-0.5930	-0.5582	-0.4203
mse	0.0126	0.0059	0.0077	0.0043	0.0225	0.0072	0.0035	0.0048	0.1049
Error Distribution: Qui-Square									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1939	0.2402	0.1485	0.2233	0.2042	0.2514	0.2086	0.2797	0.4261
mse	0.0391	0.0113	0.0266	0.0071	0.0658	0.0150	0.0047	0.0080	0.1732
$mean(\hat{\theta})$	-0.5908	-0.5744	-0.6168	-0.5836	-0.5791	-0.5679	-0.5920	-0.5572	-0.4144
mse	0.0142	0.0058	0.0087	0.0044	0.0248	0.0073	0.0035	0.0047	0.1114

Table 7: Simulation for the ARFIMA(0,0.2,1) model with different error distributions, where $\theta = 0.6$ and $n = 300$.

Error Distribution: Normal									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1253	-0.0682	0.0756	-0.0867	0.1026	-0.1083	-0.1745	-0.3861	0.1710
mse	0.0429	0.0819	0.0373	0.0886	0.0741	0.1083	0.1542	0.3562	0.0313
$mean(\hat{\theta})$	0.5241	0.3255	0.4805	0.3056	0.4873	0.2762	0.1984	-0.0753	0.5620
mse	0.0423	0.0866	0.0367	0.0936	0.0789	0.1218	0.1795	0.4784	0.0307
Error Distribution: Normal Mixture									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1318	-0.0650	0.0781	-0.0834	0.1033	-0.1071	-0.1672	-0.3839	0.1827
mse	0.0433	0.0796	0.0387	0.0861	0.0727	0.1068	0.1495	0.3535	0.0326
$mean(\hat{\theta})$	0.5286	0.3263	0.4815	0.3064	0.4873	0.2741	0.2028	-0.0780	0.5709
mse	0.0429	0.0852	0.0386	0.0924	0.0773	0.1228	0.1770	0.4826	0.0307
Error Distribution: Exponential									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1280	-0.0740	0.0777	-0.0909	0.0960	-0.1183	-0.1844	-0.3876	0.1798
mse	0.0405	0.0852	0.0367	0.0914	0.0749	0.1152	0.1627	0.3568	0.0353
$mean(\hat{\theta})$	0.5265	0.3159	0.4823	0.2981	0.4800	0.2609	0.1835	-0.0818	0.5685
mse	0.0379	0.0918	0.0354	0.0981	0.0799	0.1335	0.1929	0.4859	0.0340
Error Distribution: t-Student									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1304	-0.0707	0.0833	-0.0872	0.1070	-0.1124	-0.1796	-0.3855	0.1292
mse	0.0485	0.0840	0.0388	0.0892	0.0750	0.1109	0.1590	0.3551	0.0358
$mean(\hat{\theta})$	0.5289	0.3240	0.4896	0.3061	0.4936	0.2726	0.1922	-0.0779	0.5252
mse	0.0467	0.0880	0.0371	0.0931	0.0777	0.1251	0.1860	0.4811	0.0349
Error Distribution: Qui-Square									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1307	-0.0653	0.0780	-0.0857	0.1003	-0.1080	-0.1756	-0.3816	0.1205
mse	0.0413	0.0797	0.0385	0.0882	0.0746	0.1070	0.1556	0.3485	0.0365
$mean(\hat{\theta})$	0.5297	0.3274	0.4820	0.3051	0.4854	0.2751	0.1944	-0.0707	0.5159
mse	0.0390	0.0845	0.0372	0.0936	0.0793	0.1214	0.1839	0.4677	0.0363

The ARFIMA $(1, d, 1)$

The estimation results are presented in Tables 8-11 for $d = 0.2$, $n = 300$ and $\phi, \theta \in \{-0.2, -0.6, 0.2, 0.6\}$. For the combinations of parameters studied here the bias in d tends to be substantial with the sign of the bias determined by the sign and the relative magnitudes of the MA and AR coefficients. For $\phi < 0$ and $\theta = \pm 0.2$ the methods FT and SPER are very competitive and present better estimates of (d, ϕ, θ) . For $\theta = -0.2$, the mean squared error of GPH, ROB and SPER decreases with $g(n)$; however, the bias of the two first methods increases and the estimates also tend to underestimate the true value of d . The LOB estimator is markedly highly biased.

Tables 10-11 reveal that the bias of the estimates tends to be larger for $\phi > 0$. Now the parameter is overestimated and SPER has the best performance among all methods.

Again it appears that FT loses its superiority by presenting large bias in the estimate of the long-run and short-run parameters. Also, for this method there are parameter effects in the estimations for non-Gaussian distribution errors.

Table 8: Simulation for the ARFIMA(1,0.2,1) model with different error distributions, where $\phi = -0.6, \theta = -0.2$ and $n = 300$.

Error Distribution: Normal									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2106	0.1835	0.1587	0.1654	0.2104	0.1785	0.1487	0.0908	0.2303
mse	0.0413	0.0094	0.0267	0.0075	0.0746	0.0124	0.0087	0.0163	0.0239
$mean(\hat{\phi})$	-0.5000	-0.5811	-0.5579	-0.5980	-0.4251	-0.5804	-0.6103	-0.6623	-0.5787
mse	0.0859	0.0166	0.0501	0.0143	0.1890	0.0202	0.0154	0.0159	0.0163
$mean(\hat{\theta})$	-0.0785	-0.1863	-0.1832	-0.2208	-0.0094	-0.1913	-0.2499	-0.3603	-0.1469
mse	0.1186	0.0340	0.0760	0.0270	0.1847	0.0398	0.0282	0.0414	0.0658
Error Distribution: Normal Mixture									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1935	0.1808	0.1337	0.1576	0.1818	0.1752	0.1434	0.0969	0.2170
mse	0.0409	0.0104	0.0288	0.0089	0.0650	0.0133	0.0103	0.0151	0.0192
$mean(\hat{\phi})$	-0.4970	-0.5915	-0.5610	-0.6126	-0.4356	-0.5889	-0.6215	-0.6593	-0.5914
mse	0.1031	0.0159	0.0668	0.0144	0.1990	0.0245	0.0149	0.0162	0.0148
$mean(\hat{\theta})$	-0.0990	-0.2044	-0.2163	-0.2478	-0.0534	-0.2082	-0.2704	-0.3551	-0.1768
mse	0.1099	0.0358	0.0719	0.0306	0.1633	0.0425	0.0322	0.0406	0.0514
Error Distribution: Exponential									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2192	0.1921	0.1619	0.1694	0.2149	0.1859	0.1535	0.0875	0.2345
mse	0.0394	0.0100	0.0250	0.0075	0.0709	0.0132	0.0080	0.0171	0.0264
$mean(\hat{\phi})$	-0.5011	-0.5829	-0.5617	-0.6025	-0.4257	-0.5833	-0.6149	-0.6676	-0.5815
mse	0.0892	0.0142	0.0530	0.0124	0.1928	0.0180	0.0132	0.0156	0.0153
$mean(\hat{\theta})$	-0.0704	-0.1782	-0.1832	-0.2199	-0.0053	-0.1854	-0.2486	-0.3667	-0.1450
mse	0.1177	0.0339	0.0742	0.0264	0.1846	0.0388	0.0267	0.0435	0.0654

(continued)Table 8: Simulation for the ARFIMA(1,0.2,1) model with different error distributions, where $\phi = -0.6, \theta = -0.2$ and $n = 300$.

Error Distribution: t-Student									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2131	0.1871	0.1530	0.1655	0.2067	0.1808	0.1523	0.0859	0.2349
mse	0.0399	0.0100	0.0257	0.0074	0.0634	0.0131	0.0086	0.0176	0.0387
$mean(\hat{\phi})$	-0.4854	-0.5766	-0.5626	-0.5932	-0.4352	-0.5776	-0.6041	-0.6594	-0.5682
mse	0.0947	0.0173	0.0474	0.0156	0.1664	0.0235	0.0150	0.0160	0.0182
$mean(\hat{\theta})$	-0.0618	-0.1793	-0.1937	-0.2168	-0.0208	-0.1871	-0.2412	-0.3622	-0.1313
mse	0.1223	0.0368	0.0761	0.0281	0.1712	0.0432	0.0284	0.0432	0.0871
Error Distribution: Qui-Square									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2116	0.1842	0.1562	0.1644	0.2171	0.1810	0.1501	0.0887	0.2161
mse	0.0365	0.0094	0.0232	0.0070	0.0639	0.0118	0.0079	0.0166	0.0292
$mean(\hat{\phi})$	-0.5164	-0.5845	-0.5816	-0.6038	-0.4620	-0.5858	-0.6174	-0.6654	-0.5913
mse	0.0781	0.0157	0.0368	0.0132	0.1433	0.0204	0.0130	0.0153	0.0134
$mean(\hat{\theta})$	-0.0928	-0.1872	-0.2078	-0.2261	-0.0361	-0.1923	-0.2543	-0.3620	-0.1719
mse	0.1123	0.0341	0.0665	0.0264	0.1641	0.0408	0.0285	0.0425	0.0649

Table 9: Simulation for the ARFIMA(1,0.2,1) model with different error distributions, where $\phi = -0.6, \theta = 0.2$ and $n = 300$.

Error Distribution: Normal									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1809	0.1143	0.1383	0.0968	0.1690	0.0985	0.0721	-0.1685	0.2127
mse	0.0373	0.0170	0.0266	0.0169	0.0658	0.0230	0.0244	0.1448	0.0166
$mean(\hat{\phi})$	-0.5912	-0.6137	-0.6083	-0.6174	-0.5794	-0.6194	-0.6266	-0.7032	-0.5942
mse	0.0186	0.0048	0.0056	0.0046	0.0478	0.0053	0.0052	0.0263	0.0047
$mean(\hat{\theta})$	0.1931	0.1025	0.1354	0.0816	0.1864	0.0785	0.0451	-0.2976	0.2171
mse	0.0571	0.0270	0.0440	0.0259	0.0915	0.0384	0.0387	0.2699	0.0280
Error Distribution: Normal Mixture									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1845	0.1168	0.1314	0.0971	0.1799	0.1029	0.0751	-0.1558	0.2034
mse	0.0416	0.0173	0.0295	0.0174	0.0623	0.0221	0.0237	0.1352	0.0146
$mean(\hat{\phi})$	-0.6038	-0.6251	-0.6205	-0.6291	-0.6027	-0.6304	-0.6358	-0.7106	-0.6063
mse	0.0165	0.0050	0.0061	0.0051	0.0283	0.0056	0.0059	0.0201	0.0046
$mean(\hat{\theta})$	0.1839	0.0925	0.1155	0.0687	0.1758	0.0716	0.0379	-0.2871	0.1940
mse	0.0626	0.0292	0.0491	0.0297	0.0913	0.0392	0.0429	0.2552	0.0245
Error Distribution: Exponential									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1920	0.1120	0.1440	0.0959	0.1896	0.0969	0.0729	-0.1618	0.2074
mse	0.0372	0.0177	0.0244	0.0169	0.0669	0.0235	0.0245	0.1396	0.0153
$mean(\hat{\phi})$	-0.6019	-0.6233	-0.6143	-0.6261	-0.5979	-0.6285	-0.6338	-0.7127	-0.6024
mse	0.0124	0.0053	0.0059	0.0054	0.0324	0.0060	0.0058	0.0247	0.0049
$mean(\hat{\theta})$	0.1915	0.0869	0.1321	0.0682	0.1866	0.0646	0.0354	-0.3010	0.2004
mse	0.0594	0.0308	0.0438	0.0296	0.1005	0.0423	0.0430	0.2715	0.0267

(continued)Table 9: Simulation for the ARFIMA(1,0.2,1) model with different error distributions, where $\phi = -0.6, \theta = 0.2$ and $n = 300$.

Error Distribution: t-Student									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2039	0.1174	0.1514	0.0965	0.2056	0.1023	0.0716	-0.1627	0.1956
mse	0.0392	0.0166	0.0247	0.0172	0.0643	0.0223	0.0250	0.1403	0.0211
$mean(\hat{\phi})$	-0.6001	-0.6190	-0.6089	-0.6235	-0.5882	-0.6249	-0.6312	-0.7075	-0.6067
mse	0.0090	0.0043	0.0054	0.0043	0.0371	0.0049	0.0052	0.0241	0.0075
$mean(\hat{\theta})$	0.2058	0.0980	0.1458	0.0725	0.2129	0.0745	0.0378	-0.2975	0.1882
mse	0.0616	0.0284	0.0431	0.0287	0.0942	0.0392	0.0434	0.2677	0.0334
Error Distribution: Qui-Square									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.1918	0.1151	0.1469	0.0999	0.1918	0.1012	0.0756	-0.1640	0.1818
mse	0.0387	0.0173	0.0237	0.0163	0.0664	0.0224	0.0237	0.1409	0.0191
$mean(\hat{\phi})$	-0.6035	-0.6184	-0.6113	-0.6222	-0.5947	-0.6246	-0.6303	-0.7101	-0.6206
mse	0.0093	0.0051	0.0056	0.0049	0.0344	0.0052	0.0054	0.0191	0.0145
$mean(\hat{\theta})$	0.1930	0.0987	0.1412	0.0797	0.1955	0.0765	0.0448	-0.2977	-0.1661
mse	0.0622	0.0276	0.0401	0.0257	0.0986	0.0364	0.0390	0.2649	0.0372

Table 10: Simulation for the ARFIMA(1,0.2,1) model with different error distributions, where $\phi = 0.6, \theta = -0.2$ and $n = 300$.

Error Distribution: Normal									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2712	0.4946	0.2172	0.4746	0.3000	0.5420	0.3681	0.4323	0.6699
mse	0.0488	0.0981	0.0277	0.0829	0.0777	0.1305	0.0298	0.0543	0.3537
$mean(\hat{\phi})$	0.4960	0.2816	0.5516	0.3062	0.4624	0.2282	0.4279	0.3629	0.0163
mse	0.0553	0.1219	0.0297	0.1012	0.0900	0.1618	0.0365	0.0643	0.5492
$mean(\hat{\theta})$	-0.2264	-0.2319	-0.2213	-0.2247	-0.2372	-0.2432	-0.2029	-0.2084	-0.3576
mse	0.0113	0.0130	0.0090	0.0117	0.0137	0.0147	0.0091	0.0104	0.0960
Error Distribution: Normal Mixture									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2748	0.4948	0.2253	0.4774	0.3004	0.5407	0.3693	0.4322	0.6788
mse	0.0460	0.0972	0.0227	0.0834	0.0749	0.1295	0.0301	0.0542	0.3600
$mean(\hat{\phi})$	0.4954	0.2824	0.5483	0.3036	0.4640	0.2282	0.4262	0.3614	0.0077
mse	0.0517	0.1206	0.0261	0.1014	0.0872	0.1638	0.0371	0.0650	0.5467
$mean(\hat{\theta})$	-0.2235	-0.2310	-0.2167	-0.2250	-0.2345	-0.2442	-0.2033	-0.2097	-0.3629
mse	0.0106	0.0138	0.0085	0.0125	0.0134	0.0170	0.0099	0.0112	0.0850
Error Distribution: Exponential									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2582	0.4932	0.2169	0.4760	0.3687	0.4321	0.2722	0.5380	0.6804
mse	0.0451	0.0961	0.0245	0.0828	0.0765	0.1274	0.0298	0.0541	0.3602
$mean(\hat{\phi})$	0.5053	0.2801	0.5517	0.3003	0.4816	0.2275	0.4235	0.3587	0.0088
mse	0.0523	0.1225	0.0265	0.1042	0.0842	0.1647	0.0377	0.0657	0.5513
$mean(\hat{\theta})$	-0.2314	-0.2362	-0.2236	-0.2313	-0.2451	-0.2494	-0.2086	-0.2145	-0.3593
mse	0.0102	0.0135	0.0082	0.0122	0.0134	0.0165	0.0091	0.0105	0.0973

(continued) Table 10: Simulation for the ARFIMA(1,0.2,1) model with different error distributions, where $\phi = 0.6$, $\theta = -0.2$ and $n = 300$.

Error Distribution: t-Student									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2732	0.4992	0.2192	0.4778	0.3029	0.5468	0.3678	0.4328	0.7365
mse	0.0424	0.0987	0.0228	0.0834	0.0804	0.1328	0.0295	0.0545	0.4022
$mean(\hat{\phi})$	0.5004	0.2785	0.5535	0.3020	0.4607	0.2205	0.4271	0.3607	0.0474
mse	0.0496	0.1200	0.0247	0.1017	0.0896	0.1681	0.0360	0.0643	0.5686
$mean(\hat{\theta})$	-0.2257	-0.2354	-0.2224	-0.2308	-0.2403	-0.2510	-0.2088	-0.2149	-0.2431
mse	0.0097	0.0118	0.0078	0.0108	0.0148	0.0153	0.0081	0.0092	0.2526
Error Distribution: Qui-Square									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2714	0.4939	0.2208	0.4770	0.2878	0.5376	0.3700	0.4322	0.7620
mse	0.0496	0.0964	0.0258	0.0833	0.0826	0.1266	0.0302	0.0542	0.4210
$mean(\hat{\phi})$	0.4934	0.2780	0.5481	0.2994	0.4681	0.2249	0.4226	0.3588	0.1051
mse	0.0561	0.1241	0.0273	0.1046	0.0907	0.1665	0.0376	0.0658	0.5771
$mean(\hat{\theta})$	-0.2295	-0.2358	-0.2216	-0.2292	-0.2426	-0.2498	-0.2064	-0.2121	-0.1349
mse	0.0106	0.0142	0.0089	0.0129	0.0155	0.0171	0.0097	0.0112	0.3947

Table 11: Simulation for the ARFIMA(1,0.2,1) model with different error distributions, where $\phi = 0.6, \theta = 0.2$ and $n = 300$.

Error Distribution: Normal									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2627	0.4303	0.2137	0.4106	0.2757	0.4629	0.3389	0.3789	0.1777
mse	0.0425	0.0624	0.0244	0.0504	0.0679	0.0808	0.0211	0.0328	0.0705
$mean(\hat{\phi})$	0.4484	0.2243	0.5067	0.2521	0.4445	0.1693	0.3873	0.3340	0.5006
mse	0.0888	0.1855	0.0564	0.1615	0.1047	0.2326	0.0729	0.0981	0.1242
$mean(\hat{\theta})$	0.1108	0.0463	0.1227	0.0558	0.1170	0.0217	0.1223	0.1067	0.0795
mse	0.0406	0.0567	0.0313	0.0546	0.0481	0.0674	0.0386	0.0426	0.0731
Error Distribution: Normal Mixture									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2370	0.4172	0.1867	0.4007	0.2505	0.4521	0.3333	0.3793	0.1516
mse	0.0445	0.0577	0.0260	0.0471	0.0744	0.0765	0.0197	0.0329	0.0747
$mean(\hat{\phi})$	0.4724	0.2481	0.5318	0.2788	0.4653	0.1983	0.4014	0.3401	0.5453
mse	0.0815	0.1705	0.0473	0.1410	0.1015	0.2072	0.0637	0.0966	0.0929
$mean(\hat{\theta})$	0.1108	0.0582	0.1231	0.0741	0.1145	0.0404	0.1317	0.1137	0.1021
mse	0.0364	0.0543	0.0282	0.0494	0.0441	0.0576	0.0342	0.0430	0.0616
Error Distribution: Exponential									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2602	0.4231	0.2093	0.4078	0.2939	0.4606	0.3358	0.3797	0.1684
mse	0.0417	0.0592	0.0227	0.0493	0.0723	0.0793	0.0203	0.0330	0.0628
$mean(\hat{\phi})$	0.4464	0.2364	0.5108	0.2681	0.4214	0.1865	0.4051	0.3445	0.5110
mse	0.0918	0.1788	0.0547	0.1515	0.1181	0.2197	0.0639	0.0957	0.1092
$mean(\hat{\theta})$	0.1055	0.0501	0.1217	0.0681	0.1105	0.0351	0.1364	0.1173	0.0805
mse	0.0387	0.0589	0.0311	0.0546	0.0490	0.0685	0.0362	0.0441	0.0596

(continued) Table 11: Simulation for the ARFIMA(1,0.2,1) model with different error distributions, where $\phi = 0.6, \theta = 0.2$ and $n = 300$.

Error Distribution: t-Student									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2655	0.4288	0.2142	0.4100	0.2798	0.4610	0.3369	0.3796	0.0818
mse	0.0453	0.0620	0.0246	0.0502	0.0740	0.0806	0.0208	0.0330	0.1034
$mean(\hat{\phi})$	0.4518	0.2232	0.5065	0.2566	0.4266	0.1717	0.3974	0.3387	0.5880
mse	0.0866	0.1933	0.0593	0.1622	0.1217	0.2384	0.0697	0.1002	0.1049
$mean(\hat{\theta})$	0.1192	0.0459	0.1259	0.0622	0.1050	0.0238	0.1330	0.1144	0.0760
mse	0.0394	0.0650	0.0323	0.0596	0.0520	0.0730	0.0385	0.0454	0.0728
Error Distribution: Qui-Square									
Estimate	GPH($g(n)$)		SPER ($g(n)$)		ROB($g(n)$)		LOB(τ)		FT
$g(n)$ or τ	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.7}$	$n^{0.8}$	-
$mean(\hat{d})$	0.2555	0.4217	0.2066	0.4027	0.2694	0.4544	0.3332	0.3801	0.0518
mse	0.0422	0.0599	0.0236	0.0479	0.0704	0.0791	0.0198	0.0331	0.1110
$mean(\hat{\phi})$	0.4466	0.2411	0.5177	0.2800	0.4401	0.1841	0.4082	0.3484	0.6137
mse	0.0932	0.1796	0.0483	0.1394	0.1064	0.2295	0.0585	0.0888	0.0946
$mean(\hat{\theta})$	0.0958	0.0481	0.1212	0.0697	0.1008	0.0215	0.1311	0.1153	0.0644
mse	0.0430	0.0621	0.0307	0.0519	0.0473	0.0728	0.0341	0.0410	0.0732

3.3 Model misspecification.

As demonstrated in the previous section the FT method is superior than the others semiparametric methods in the case of ARFIMA(0, d , 0) models. However, this estimator has showed large bias depending on the sign and on the relative magnitude of the AR or MA components. In this section, we investigate the bias of the estimates in the FT method when the model order is misspecified. We simulated ARFIMA(0, d , 0) process and indicated for the method an ARFIMA(1, d , 0) and ARFIMA(0, d , 1) models. Also, ARFIMA(1, d , 0) was simulated and the specified order was ARFIMA(0, d , 0). Note that the semiparametric methods do not depend on the corrected specification of the spectral density function to obtain the estimate of the fractional parameter. Hence, the bias of d in the semiparametric estimates is unaffected by the precise order of the ARMA process. However, a misspecification of the FT procedure may lead to bias in both short-run and long-run parameter estimates.

In addition to the semiparametric results, the tables also give the FT estimates for both the wrong and the correct models. In Table 12 are the results when the generating process in an ARFIMA(0, d , 0) for $d = 0.2$ and 0.4 . In this case, the correlation structure of the series is only associated with d . The difference between the FT estimates for the correct and the uncorrect orders is around -0.06 (see the second part of Table 12).

Considering an AR or MA order for the FT method, the bias of the long memory parameter increases substantially as d gets close to non-stationary condition. In Table 13 we have the situation where the true model is an ARFIMA(1, d , 0). The method FT presents a markedly higher bias of d if the selected process is an ARFIMA(0, d , 0), since the correlation structure of the short-run and long-run parameters is only retained by \hat{d} . Note that this is more evidente as ϕ increases.

Table 12: Estimation results when generating process is an ARFIMA (0, d, 0)

(d=0.2)						
Method	GPH(g(n))		SPER (g(n))		ROB(g(n))	
	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$
$g(n)$ or τ						
$mean(\hat{d})$	0.2022	0.2047	0.1519	0.1839	0.2007	0.2047
$bias(\hat{d})$	0.0022	0.0047	-0.0482	-0.1612	0.0006	0.0047
$mse(\hat{d})$	0.0423	0.0105	0.0262	0.0071	0.0681	0.0129
<i>mean</i>	-	-	-	-	-	-
<i>mse</i>	-	-	-	-	-	-
Method	LOB(τ)		FT (AR)	FT (MA)	FT (*)	
	$n^{0.7}$	$n^{0.8}$	$p=1, q=0$	$p=0, q=1$	$p=q=0$	
$g(n)$ or τ						
$mean(\hat{d})$	0.1726	0.1675	0.1904	0.2021	0.1981	
$bias(\hat{d})$	-0.0275	-0.0326	-0.0097	0.0021	-0.0019	
$mse(\hat{d})$	0.0068	0.0045	0.0089	0.0075	0.0023	
<i>mean</i>	-	-	$\hat{\phi}=0.0086$	$\hat{\theta}=0.0037$	-	
<i>mse</i>	-	-	0.0116	0.0104	-	
(d=0.4)						
Method	GPH(g(n))		SPER (g(n))		ROB(g(n))	
	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$	$n^{0.5}$	$n^{0.7}$
$g(n)$ or τ						
$mean(\hat{d})$	0.4112	0.4068	0.3515	0.3876	0.4049	0.4044
$bias(\hat{d})$	0.0112	0.0067	-0.0485	-0.0125	-0.0049	0.0044
$mse(\hat{d})$	0.0401	0.0099	0.0265	0.0071	0.0662	0.0126
<i>mean</i>	-	-	-	-	-	-
<i>mse</i>	-	-	-	-	-	-
Method	LOB(τ)		FT (AR)	FT (MA)	FT (*)	
	$n^{0.7}$	$n^{0.8}$	$p=1, q=0$	$p=0, q=1$	$p=q=0$	
$g(n)$ or τ						
$mean(\hat{d})$	0.3192	0.3239	0.4518	0.4693	0.4020	
$bias(\hat{d})$	-0.0808	-0.0762	0.0517	0.0692	0.0020	
$mse(\hat{d})$	0.0095	0.0075	0.0208	0.0248	0.0122	
<i>mean</i>	-	-	$\hat{\phi}=0.0407$	$\hat{\theta}=0.0706$	-	
<i>mse</i>	-	-	0.0155	0.0295	-	

Notes: *means the process with the corrected order; n=300

Table 13: Estimation results when generating process is an ARFIMA(1, d, 0)

(d=0.2, φ=0.2)						
Method	GPH(g(n))		SPER (g(n))		ROB(g(n))	
	<i>n</i> ^{0.5}	<i>n</i> ^{0.7}	<i>n</i> ^{0.5}	<i>n</i> ^{0.7}	<i>n</i> ^{0.5}	<i>n</i> ^{0.7}
<i>g</i> (n) or τ						
<i>mean</i> (\hat{d})	0.2135	0.2534	0.1659	0.2341	0.2204	0.2625
<i>bias</i> (\hat{d})	0.0134	0.0534	-0.0340	0.0340	0.0204	0.0624
<i>mse</i> (\hat{d})	0.0393	0.0137	0.0252	0.0079	0.0652	0.0175
<i>mean</i> ($\hat{\phi}$)	0.1916	0.1406	0.2339	0.1573	0.1967	0.1334
<i>mse</i> ($\hat{\phi}$)	0.0393	0.0152	0.0283	0.0099	0.0623	0.0186
Method	LOB(τ)		FT	FT (*)		
<i>g</i> (n) or τ	<i>n</i> ^{0.7}	<i>n</i> ^{0.8}	<i>p</i> = <i>q</i> =0	<i>p</i> =1, <i>q</i> =0		
<i>mean</i> (\hat{d})	0.2172	0.2571	0.3594	0.1762		
<i>bias</i> (\hat{d})	0.1772	0.0570	0.1593	-0.0238		
<i>mse</i> (\hat{d})	0.0051	0.0055	0.0377	0.0194		
<i>mean</i> ($\hat{\phi}$)	0.1748	0.1359	-	0.2138		
<i>mse</i> ($\hat{\phi}$)	0.0069	0.0067	-	0.0221		
(d=0.0.2, φ=0.6)						
Method	GPH(g(n))		SPER (g(n))		ROB(g(n))	
	<i>n</i> ^{0.5}	<i>n</i> ^{0.7}	<i>n</i> ^{0.5}	<i>n</i> ^{0.7}	<i>n</i> ^{0.5}	<i>n</i> ^{0.7}
<i>g</i> (n) or τ						
<i>mean</i> (\hat{d})	0.2532	0.4661	0.2108	0.4502	0.2706	0.5078
<i>bias</i> (\hat{d})	-0.0532	0.2661	0.0108	0.2502	-0.0706	0.3077
<i>mse</i> (\hat{d})	0.0408	0.0800	0.0222	0.0684	0.0631	0.1066
<i>mean</i> ($\hat{\phi}$)	0.5171	0.3279	0.5618	0.3479	0.5029	2919
<i>mse</i> ($\hat{\phi}$)	0.0404	0.0839	0.0216	0.0709	0.0671	0.1074
Method	LOB(τ)		FT	FT (*)		
<i>g</i> (n) or τ	<i>n</i> ^{0.7}	<i>n</i> ^{0.8}	<i>p</i> = <i>q</i> =0	<i>p</i> =1, <i>q</i> =0		
<i>mean</i> (\hat{d})	0.3586	0.4143	0.7363	0.4957		
<i>bias</i> (\hat{d})	0.1586	0.2132	0.5364	-0.2957		
<i>mse</i> (\hat{d})	0.0265	0.0459	0.2911	0.2211		
<i>mean</i> ($\hat{\phi}$)	0.4376	0.3846	-	0.3272		
<i>mse</i> ($\hat{\phi}$)	0.0297	0.0488	-	0.1759		

Notes: *means the process with the corrected order; n=300

4. Concluding Remarks.

We investigated how some parametric and nonparametric estimation methods of the long memory parameter in the ARFIMA (p, d, q) model can be affected when the innovations errors are from heavy-tailed, skewed and bimodal distributions. The semiparametric methods do not present any evidence against the violation of the Gaussian assumption. However, the FT method seems to be sensitive when t-Student and Qui-square distributions are considered for the errors innovations. In these situations there are large bias of the parameters and the behaviour of the estimates depend on the magnitude of the short-run components. The bias in the FT estimator was also investigated for the uncorrected ARFIMA specification. The effect of small and large sample sizes was also considered. The results reported in this paper increase our understanding of the finite sample properties of the estimators considered.

Acknowledgements

V.A. Reisen was partially supported by CNPq-Brazil. S.R.C. Lopes was partially supported by Pronex *Fenômenos Críticos em Probabilidade e Processos Estocásticos* (Convênio FINEP/MCT/CNPq 41/ 96/0923/00) and CNPq-Brazil.

The authors would like to thank Professor Francisco Cribari-Neto (UFPE) and an anonymous referee for some suggestions that improved this paper. V. A. Reisen thanks his undergraduate student Artur J. Lemonte (PIBIC-UFES) for helping with some simulations presented in Section 3.3.

Submitted in February 2001. Revised in April 2002.

References

- Anderson, T.W. & A. M. Walker 1964. “On the asymptotic distribution of the autocorrelations of a sample from a linear stochastic process”. *Annals of Mathematical Statistics* **35**:1296–1303.
- Beran, J. 1994. “*Statistics for long memory processes*. Chapman and Hall: New York.
- Bisaglia, L. & D. Guéran 1998. “A comparison of techniques of estimation in long memory processes”. *Computational Statistics and Data Analysis* **27**:61–81.
- Cheung, Y-W & F.X. Diebold 1994. “On maximum-likelihood estimation of the differencing parameter of fractionally integrated noise with unknown mean”. *Journal of Econometrics* **62**:301–316.
- Dahlhaus, R. 1989. “Efficient parameter estimation for self-similar processes”. *The Annals of Statistics* **17**(4):1749–1766.
- Fox, R. & M.S. Taqqu 1986. “Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series”. *The Annals of Statistics* **14**(2):517–532.
- Geweke, J. & S. Porter-Hudak 1983. “The estimation and application of long memory time series model”. *Journal of Time Series Analysis* **4**(4):221–238.
- Giraitis, L. & M.S. Taqqu 1999. “Whittle estimator for finite-variance non-Gaussian time series with long memory”. *The Annals of Statistics* **27**(1):178–203.
- Hosking, J. 1981. “Fractional differencing”. *Biometrika* **68**(1):165–176.
- Hosking, J. 1984. “Modelling persistence in hydrological time series using fractional differencing”. *Water Resources Research* **20**(12):1898–1908.

- Huber, P. 1996. "Robust statistical procedures". CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, 68, 2nd edition.
- Hurvich, C.M., R.S. Deo & J. Brodsky 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* **19**(1):19–46.
- Hurvich, C.M. & R.S. Deo 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* **20**(3):331–341.
- Little, R.J.A. & D.B. Rubin 1987. *Statistical analysis with missing data*. John Wiley: New York.
- Ljung, G.M. & G.E.P. Box 1978. "On a measure of lack of fit in time series models". *Biometrika* **65**(2):297–303.
- Lobato, I. & P.M. Robinson 1996. "Averaged periodogram estimation of long memory". *Journal of Econometrics* **73**:303–324.
- Reisen, V.A. 1994. "Estimation of the fractional difference parameter in the ARFIMA(p, d, q) model using the smoothed periodogram". *Journal of Time Series Analysis* **15**(3):335–350.
- Reisen, V.A. & S.R.C. Lopes 1999. "Some simulations and applications of forecasting long-memory time series models". *Journal of Statistical Planning and Inference* **80**(2):269–287.
- Reisen, V.A., B.Abraham & E.M.M. Toscano 2000. "Parametric and semiparametric estimations of stationary univariate ARFIMA models". *Brazilian Journal of Probability and Statistics* **14**:185–206.
- Reisen, V.A., B.Abraham & S.R.C. Lopes 2001. "Estimation of Parameters in ARFIMA Processes: A simulation study. *Communications in Statistics Simulation and Computation* **30** (4):787–803.

- Robinson, P.M. 1994. "Semiparametric analysis of long-memory time series". *The Annals of Statistics* **22**(1):515–539.
- Robinson, P.M. 1995. "Log-periodogram regression of time series with long range dependence". *The Annals of Statistics* **23**(3):1048–1072.
- Smith, J., N. Taylor & S. Yadav 1997. "Comparing the bias and misspecification in ARFIMA models". *Journal of Time Series Analysis* **18**(5):507–527.
- Sowell, F. 1992. "Maximum likelihood estimation of stationary univariate fractionally integrated time series models". *Journal of Econometric* **53**:165–188.
- Velasco, C. 2000. "Non-Gaussian Log-periodogram regression". *Econometric Theory* **16**:44–79.
- Whittle, P. 1953. "Estimation and information in stationary time series". *Ark. för Matematik* **2**:423–434.