E C O N O M I C S B U L L E T I N

The exact maximum likelihood estimation of ARFIMA processes and model selection criteria: A Monte Carlo study

Sandrine Lardic
University of Paris 10 − MODEM

Valerie Mignon
University of Paris 10 − THEMA

Abstract

We propose a detailed Monte Carlo study of model selection criteria when the exact maximum likelihood (EML) method is used to estimate ARFIMA processes. More specifically, our object is to assess the performance of two automatic selection criteria in the presence of long−term memory: Akaike and Schwarz information criteria. Two special processes are considered: a pure fractional noise model (ARFIMA(0,d,0)) and an ARFIMA(1,d,0) process. For each criterion, we compute bias and root mean squared error for various d and AR(1) parameter values. Obtained results suggest that the Schwarz information criterion frequently selects the right model. Moreover, this criterion outperforms the other one in terms of bias and RMSE, for both pure fractional noise and ARFIMA processes.

We are grateful to Emmanuel Dubois for helpful comments and invaluable computer assitance. We also thank an anonymous referee for his comments.
Citation: Lardic, Sandrine and Valerie Mignon, (2004) "The exact maximum likelihood estimation of ARFIMA processes and model selection criteria: A Monte Carlo study." *Economics Bulletin,* Vol. 3, No. 21 pp. 1−16
Submitted: May 6, 2004. **Accepted:** June 19, 2004.
URL: http://www.economicsbulletin.com/2004/volume3/EB−04C20020A.pdf

1 Introduction

A vast literature has been published on long-term memory processes, both from theoretical and empirical viewpoints[1]. Such processes can be characterized by an autocorrelation function which decreases at an hyperbolic rate. In other words, in the frequency domain, the spectral density exhibits a pole at zero frequency. Autoregressive fractionally integrated moving average (ARFIMA) processes were introduced by Granger and Joyeux (1980) and Hosking (1981) to account for these two characteristics. Roughly speaking, these processes are characterized by a fractional differencing parameter d which accounts for the long-term dynamics, while traditional AR and MA components capture the short-term dynamics of the time series.

The growing empirical literature on ARFIMA processes highlights the importance of efficient estimation procedures for these processes. Several estimation techniques have been proposed. More specifically, one can distinguish one-step and two-step estimation methods (see e.g. Sowell (1992) for details). Concerning two-step procedures[2], one estimates, in the first step, the fractional differencing parameter. AR and MA parameters are then estimated in the second step. The drawback with these methods is that they use information only at low frequencies. In other words, they do not take account of the short-term properties of the series when estimating the fractional differencing parameter. This is an important issue since the estimate of the long-term parameter could be contaminated by the presence of short-term components. One-step procedures in which all parameters of the $ARFIMA(p, d, q)$ representation are estimated simultaneously have been suggested to avoid this problem. In these procedures, information at high frequencies is also used. Among these methods, one can distinguish exact maximum likelihood (Dahlhaus (1989), An *et al.* (1992) and Sowell (1992)) and approximate maximum likelihood (Li and McLeod (1986) and Fox and Taqqu (1986))[3]. We propose here to focus on the exact maximum likelihood method (EML) developed by Sowell (1992). This topic is relevant since, according to Dahlhaus (1988) and Sowell (1992), EML is the most efficient estimation procedure for ARFIMA processes (see also Cheung and Diebold (1994)). It permits us to tackle the short memory contamination problem and the small-sample bias associated notably with the popular Geweke and Porter-Hudak two-step method. However, despite these advantages, this technique has received little attention in the literature, compared to the Geweke and Porter-Hudak procedure, because the EML estimation of a long-range dependent process is computationally intensive for high dimensional systems, due to the time required to compute the elements of the covariance matrix of the process and its inverse (see, for example, Bollerslev and Jubinski (1999))[4]. This procedure has thus been criticized as too computationally demanding, while the afore-mentioned methods have been criticized as inaccurate for finite samples (see e.g. Sowell (1992)). It constitutes however the preferable estimation procedure, especially in small and medium sample sizes.

We present here a Monte Carlo study based on two special processes: $ARFIMA(0, d, 0)$ and $ARFIMA(1, d, 0)$ processes. $ARFIMA(0, d, 0)$ corresponds to a pure fractional noise, *i.e.* a process which exhibits long-range dependence only (if $d \neq 0$). $ARFIMA(1, d, 0)$ is a process with both short-term and long-term components[5]. More specifically, we study the ability of the EML method to detect the right process according to two criteria: Akaike information and Schwarz information criteria[6]. For each criterion, we also compute bias and root mean squared error (RMSE) in association with various d and $AR(1)$ parameter values. To our knowledge, no result is available regarding the success of different model selection criteria when EML is used to estimate

[1] For a survey, the reader is referred to Baillie (1996), Lardic and Mignon (1999) and Robinson (2003) among others.

[2] The most popular is the technique developed by Geweke and Porter-Hudak (1983). Other methods have also been proposed by Janacek (1982) and Shea (1991). See also Künsch (1987) and Robinson (1995) who introduced the Gaussian semiparametric method.

[3] Note that some refined semiparametric methods have also been suggested (see Andrews and Guggenberger (2003) and Andrews and Sun (2004)).

[4] For a recent discussion on this subject, see Doornik and Ooms (2003).

[5] It is well known that the combination of a large AR component with a long memory component is hard to estimate (see, for example, Agiakloglou *et al.* (1993)).

[6] See Beran, Bhansali and Ocker (1998) for a discussion of these criteria for $ARFIMA(p, d, 0)$ models.

ARFIMA processes[7].

The paper will proceed as follows. In section 2, we recall some definitions of ARFIMA processes and some details concerning the EML method. Section 3 describes results from Monte Carlo simulations, for both $ARFIMA(0,d,0)$ and $ARFIMA(1,d,0)$ processes. Section 4 concludes.

2 Some definitions

A time series X_t, $t = 1, ..., T$, follows an $ARFIMA(p,d,q)$ process if:

$$\Phi(L)(1-L)^d X_t = \Theta(L)\varepsilon_t \tag{1}$$

where $\Phi(L) = 1 - \phi_1 L - ... - \phi_p L^p$, $\Theta(L) = 1 - \theta_1 L - ... - \theta_q L^q$, $\varepsilon_t \sim iid(0, \sigma^2)$, and:

$$(1-L)^d = 1 - dL - \frac{d(1-d)}{2!}L^2 - \frac{d(1-d)(2-d)}{3!}L^3 - .... = \sum_{j=0}^{\infty} \pi_j L^j \tag{2}$$

with:

$$\pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}, \tag{3}$$

Γ being the gamma function.

X_t, $t = 1, ..., T$, is both stationary and invertible if the roots of $\Phi(L)$ and $\Theta(L)$ are outside the unit circle and $-1/2 < d < 1/2$.

The parameters of this $ARFIMA(p,d,q)$ process can be jointly estimated by the EML method, assuming $\varepsilon_t \sim iidN(0, \sigma^2)$ (see Sowell (1992) for details). Let $X_t, t = 1, ..., T$, be a fractionally integrated stationary Gaussian time series. $X = (X_1, ..., X_T)'$ follows a normal law with mean zero and covariance matrix Σ. Its density function is given by:

$$f(X, \Sigma) = (2\pi)^{-T/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2} X_t' \Sigma^{-1} X_t\right) \tag{4}$$

Due to the stationarity property, the covariance matrix has a Toeplitz form: $\Sigma = [\gamma_{i-j}]$ with $i, j = 1, 2, ..., T$. Estimation of ARFIMA processes by EML requires writing the spectral density function of X_t, denoted as $f_X(\lambda)$, in terms of the parameters of the model and then evaluating the autocovariance function γ_s at lag s by:

$$\gamma_s = \frac{1}{2\pi} \int_0^{2\pi} f_X(\lambda) e^{i\lambda s} d\lambda \tag{5}$$

The EML estimator of d has an asymptotic normal distribution (see Dahlhaus (1989)). The pertinence of this method lies in its using all information concerning the short and long-term behavior of the series since it estimates all parameters of the $ARFIMA(p,d,q)$ representation simultaneously.

[7] Only a few Monte Carlo studies have been implemented to assess the performance of automatic selection criteria in the presence of long memory and these do not concern the EML method. Schmidt and Tschernig (1993) studied the performance of selection criteria in the single case of pure fractionally integrated processes and for the Whittle-type approximate maximum likelihood method. Crato and Ray (1996) extended this study by considering more long-term memory processes and by studying three estimation procedures: Geweke and Porter-Hudak (1983)'s spectral regression method, Haslett and Raftery (1989)'s approximate time domain maximum likelihood procedure and Fox and Taqqu (1986)'s approximate frequency domain maximum likelihood technique.

Table 1: Number of ARFIMA(0,d,0) models selected by the AIC criterion

	−0.5	−0.4	−0.3	−0.2	−0.1	0	0.1	0.2	0.3	0.4	0.5
$(0,d,0)$	200	89	216	**246**	104	**245**	**329**	**310**	**294**	**276**	183
$(0,d,1)$	66	86	49	39	38	30	25	25	23	19	29
$(0,d,2)$	34	42	33	27	21	20	19	22	19	17	21
$(0,d,3)$	17	20	16	15	16	16	16	14	14	13	8
$(1,d,0)$	104	154	98	99	164	89	51	49	50	55	99
$(1,d,1)$	53	52	40	45	56	54	54	57	60	57	89
$(1,d,2)$	6	5	7	6	6	4	4	3	6	10	20
$(1,d,3)$	10	11	9	11	11	8	9	13	9	12	11
$(2,d,0)$	39	59	56	59	70	62	55	58	76	79	99
$(2,d,1)$	19	21	19	20	19	22	21	23	28	23	22
$(2,d,2)$	**228**	**244**	**220**	203	**213**	193	174	154	140	128	79
$(2,d,3)$	36	32	32	34	36	33	26	31	21	18	24
$(3,d,0)$	18	20	28	22	36	32	31	33	37	35	38
$(3,d,1)$	10	12	16	16	14	11	17	22	24	25	23
$(3,d,2)$	96	98	99	89	119	109	101	122	137	170	**185**
$(3,d,3)$	64	55	62	69	77	72	68	64	62	63	70

Maximum in bold.

3 Monte Carlo simulations

Two types of processes are simulated: An $ARFIMA(0,d,0)$ model, *i.e.* a model with only a long-term component, and an $ARFIMA(1,d,0)$ model, *i.e.* a model including both long-term and short-term components. For each d value ($-0.5 \leq d \leq 0.5$), we generate 1000 series of 300 observations, a size that is common with business and economic data. We run the exact maximum likelihood procedure on these series and estimate 16 $ARFIMA(p,d,q)$ processes: from $(0,d,0)$ to $(3,d,3)$ process. Then, we retain the model which maximizes Akaike (AIC) and Schwarz (SIC) information criteria.

3.1 Pure fractionally integrated processes: $ARFIMA(0,d,0)$

We first simulate $ARFIMA(0,d,0)$ processes in order to assess the performance of selection criteria in the absence of a short-term component.

3.1.1 General comments

Tables 1 and 2 give the number of $ARFIMA(0,d,0)$ processes selected by AIC and SIC respectively. The SIC criterion performs very well for positive d values since it very often retains the right model (the percentage is greater than 90%, except when $d = 0.5$). For negative d values, except when $d = -0.4$, SIC tends to select the right model, but the percentages are not very high. The results in table 2 also indicate that the second model which is relatively frequently chosen by SIC is $ARFIMA(1,d,0)$. Similar comments can be made when using the results reported in table 1. The most chosen model by AIC is $ARFIMA(0,d,0)$ for positive d values. Note however that percentages in selecting the right model are much lower than those issued from SIC criterion. For negative d values, except when $d = -0.2$, AIC tends to select the $ARFIMA(2,d,2)$ process. According to both information criteria, the most difficult process to be identified is $ARFIMA(0,-0.4,0)$, which means that the anti-persistent fractional noise is harder to identify than the persistent one. Globally, fractional noise models with positive d values are easier to identify than anti-persistent ones.

Table 2: Number of ARFIMA(0,d,0) models selected by the SIC criterion

	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5
$(0,d,0)$	**493**	275	**565**	**640**	**523**	**901**	**943**	**943**	**934**	**917**	**833**
$(0,d,1)$	198	257	155	72	84	53	19	15	15	17	18
$(0,d,2)$	6	12	3	5	5	3	3	3	3	3	2
$(0,d,3)$	1	1	1								
$(1,d,0)$	283	**428**	251	255	356	56	24	28	34	47	109
$(1,d,1)$	7	9	5	6	8	2	1	1	2	5	9
$(1,d,2)$											1
$(1,d,3)$											
$(2,d,0)$	6	14	16	15	18	9	5	6	7	6	21
$(2,d,1)$	1	2	1	2	2	2	1	1	2	1	1
$(2,d,2)$	5	2	2	4	3	3	2	1		1	3
$(2,d,3)$											
$(3,d,0)$			1	1	1	1	2	2	3	3	3
$(3,d,1)$											
$(3,d,2)$											
$(3,d,3)$											

Maximum in bold.

3.1.2 Bias and RMSE

In order to complete these results, tables 3 and 4 give the bias and root mean squared error (RMSE) calculated for each criterion. Three biases and RMSE are given:

- Case (1): bias and RMSE for the model selected by the considered criterion,
- Case (2): bias and RMSE conditional on the right model being selected by the considered criterion,
- Case (3): bias and RMSE for the right model (even if this model had not been selected by the considered criterion).

According to tables 3 and 4, and for positive d values, one can remark that bias and RMSE are very close in cases (2) and (3) for the SIC criterion. This result is not surprising since this criterion generally selects the right model for positive d values (see table 2).

According to table 3, and when the right model is selected (case (2)), SIC gives the lowest bias for negative d values (except for $d = -0.1$) and so does AIC for positive d values. Let us now consider the model selected by each criterion ($ARFIMA(0, d, 0)$ or not). In this case (case (1)), SIC always leads to the lowest bias. Note that one reason is that SIC selects the right model more often. From a general viewpoint, one can also note that, whatever the criterion, the bias is generally negative; *i.e.* d tends to be underestimated.

Let us now comment on the results reported in table 4. In case (2), AIC leads to the lowest RMSE, except for four values of d (-0.4, -0.3, 0, and 0.5), for which RMSE is minimized by SIC. In case (1), the lowest RMSE is again always given by SIC.

Finally, the study of bias and RMSE according to d values leads to the two following comments. First, concerning AIC, one can note that bias and RMSE increase along with the value of d in case (1). The same comment can be made for case (2), but only when d is positive. For negative d values, the bias is generally small and tends to decrease as d increases (notably until $d = -0.3$). In case (2) for AIC, one remarks that RMSE is relatively stable for positive d values (except for $d = 0.5$). Second, the bias associated with SIC criterion for case (1) tends to increase with the value of d, when d is positive. The bias is relatively stable for negative d values (except for $d = -0.2$). For the same case, RMSE is not very sensitive to d, but tends to increase when d tends

Table 3: ARFIMA(0,d,0). Bias

	AIC		SIC		
d	(1)	(2)	(1)	(2)	(3)
-0.5	$-6.57.10^{-2}$	$-2.80.10^{-2}$	*$-2.45.10^{-2}$*	**$-2.57.10^{-2}$**	$2.31.10^{-1}$
-0.4	$-9.93.10^{-2}$	$-1.54.10^{-2}$	*$-2.92.10^{-2}$*	**$-1.28.10^{-2}$**	$2.83.10^{-1}$
-0.3	$-1.15.10^{-1}$	6.0810^{-3}	*$-2.33.10^{-2}$*	**$4.43.10^{-3}$**	$1.27.10^{-1}$
-0.2	$-1.32.10^{-1}$	$-2.81.10^{-2}$	*$-4.21.10^{-2}$*	**$-2.69.10^{-2}$**	$5.24.10^{-2}$
-0.1	$-1.54.10^{-1}$	**$-2.36.10^{-3}$**	*$-2.53.10^{-2}$*	$3.81.10^{-2}$	$7.10.10^{-2}$
0	$-1.77.10^{-1}$	$1.10.10^{-2}$	*$-2.16.10^{-2}$*	**$7.12.10^{-3}$**	$7.20.10^{-3}$
0.1	$-1.92.10^{-1}$	**$-1.43.10^{-2}$**	*$-3.50.10^{-2}$*	$-1.68.10^{-2}$	$-1.66.10^{-2}$
0.2	$-2.39.10^{-1}$	**$-1.61.10^{-2}$**	*$-4.07.10^{-2}$*	$-1.89.10^{-2}$	$-1.86.10^{-2}$
0.3	$-2.95.10^{-1}$	**$-1.85.10^{-2}$**	*$-5.08.10^{-2}$*	**$-2.13.10^{-2}$**	$-2.05.10^{-2}$
0.4	$-3.55.10^{-1}$	**$-2.43.10^{-2}$**	*$-6.80.10^{-2}$*	$-2.72.10^{-2}$	$-2.58.10^{-2}$
0.5	$-5.29.10^{-1}$	$-6.34.10^{-2}$	*$-1.59.10^{-1}$*	**$-5.90.10^{-2}$**	$-5.59.10^{-2}$

(1): Bias given by the selected model. (2): Bias given by the selected right model. (3): Bias corresponding to the right model. In italics: minimum bias according to (1). In bold: minimum bias according to (2).

Table 4: ARFIMA(0,d,0). RMSE

	AIC		SIC		
d	(1)	(2)	(1)	(2)	(3)
-0.5	$2.55.10^{-1}$	**$5.05.10^{-2}$**	*$1.18.10^{-1}$*	$5.19.10^{-2}$	$3.55.10^{-1}$
-0.4	$2.82.10^{-1}$	$6.93.10^{-2}$	*$1.35.10^{-1}$*	**$6.54.10^{-2}$**	$3.45.10^{-1}$
-0.3	$2.90.10^{-1}$	$4.61.10^{-2}$	*$1.34.10^{-1}$*	**$4.49.10^{-2}$**	$2.00.10^{-1}$
-0.2	$2.95.10^{-1}$	**$4.54.10^{-2}$**	*$1.32.10^{-1}$*	$5.19.10^{-2}$	$1.29.10^{-1}$
-0.1	$3.19.10^{-1}$	**$7.23.10^{-2}$**	*$1.50.10^{-1}$*	$8.81.10^{-2}$	$9.85.10^{-2}$
0	$3.49.10^{-1}$	$3.54.10^{-2}$	*$1.25.10^{-1}$*	**$3.53.10^{-2}$**	$3.46.10^{-2}$
0.1	$3.67.10^{-1}$	**$4.51.10^{-2}$**	*$1.23.10^{-1}$*	$4.78.10^{-2}$	$4.76.10^{-2}$
0.2	$4.23.10^{-1}$	**$4.73.10^{-2}$**	*$1.35.10^{-1}$*	$5.07.10^{-2}$	$5.04.10^{-2}$
0.3	$4.83.10^{-1}$	**$4.66.10^{-2}$**	*$1.56.10^{-1}$*	$5.07.10^{-2}$	$5.04.10^{-2}$
0.4	$5.42.10^{-1}$	**$4.75.10^{-2}$**	*$1.86.10^{-1}$*	$5.13.10^{-2}$	$5.07.10^{-2}$
0.5	$6.75.10^{-1}$	$6.89.10^{-2}$	*$3.04.10^{-1}$*	**$6.57.10^{-2}$**	$6.29.10^{-2}$

(1): RMSE given by the selected model. (2): RMSE given by the selected right model. (3): RMSE corresponding to the right model. In italics: minimum RMSE according to (1). In bold: minimum RMSE according to (2).

Table 5: Number of ARFIMA(1,d,0) processes selected by the AIC criterion

	d										
ϕ	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5
-0.9	**349**	**437**	**536**	**464**	**389**	**376**	**410**	**457**	**397**	230	157
-0.7	**393**	**386**	**399**	**453**	**421**	**421**	**394**	**279**	**322**	**288**	**267**
-0.5	**401**	**378**	**368**	**376**	**374**	**375**	**385**	**421**	**343**	**282**	**345**
-0.3	**310**	**347**	**337**	**311**	**318**	**318**	**325**	**346**	**410**	**330**	**252**
-0.1	**186**	153	146	**239**	**186**	140	139	137	134	**190**	**305**
0.1	104	159	195	111	147	176	77	52	59	51	27
0.3	**365**	**363**	**339**	**345**	**343**	**315**	**296**	**262**	**260**	**230**	**147**
0.5	**531**	**462**	**420**	**398**	**398**	**387**	**383**	**348**	**329**	**313**	**221**
0.7	**595**	**533**	**529**	**568**	**468**	**456**	**506**	**400**	**373**	**335**	**238**
0.9	**545**	**568**	**608**	**590**	**563**	**546**	**511**	**514**	**518**	**438**	200

In bold: model most often selected.

Table 6: Number of ARFIMA(1,d,0) processes selected by the SIC criterion

	d										
ϕ	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5
-0.9	**670**	**781**	**858**	**794**	**700**	**702**	**760**	**902**	**828**	**598**	413
-0.7	**900**	**900**	**893**	**902**	**897**	**907**	**826**	**701**	**766**	**832**	**766**
-0.5	**850**	**817**	**815**	**826**	**839**	**838**	**849**	**848**	**721**	**640**	**794**
-0.3	**553**	**726**	**695**	**567**	**569**	**590**	**605**	**652**	**715**	**619**	**558**
-0.1	**415**	294	256	**500**	197	132	138	145	159	234	**507**
0.1	129	367	**563**	329	340	265	63	50	49	41	30
0.3	**697**	**696**	**723**	**800**	**817**	**736**	**758**	**677**	**651**	**653**	**622**
0.5	**944**	**931**	**929**	**923**	**917**	**913**	**908**	**902**	**897**	**881**	**828**
0.7	**970**	**968**	**965**	**955**	**951**	**943**	**940**	**930**	**925**	**902**	**855**
0.9	**922**	**954**	**974**	**970**	**976**	**969**	**962**	**953**	**947**	**914**	**747**

In bold: model most often selected.

to 0.5. Regarding now case (2), bias and RMSE increase with d, when d is positive. When d is negative, the bias is very small and tends to decrease as d increases (notably until $d = -0.3$).

To sum up, our simulations indicate that the Sowell (1992)'s procedure performs well in estimating $ARFIMA(0, d, 0)$ if the model is selected by information criteria, more especially by SIC which often retains the right model. This conclusion relating to the success of the SIC criterion confirms the results obtained by Schmidt and Tschernig (1993) and Crato and Ray (1996) with other estimation procedures (see footnote 7).

3.2 Introduction of a short-term component: $ARFIMA(1, d, 0)$ processes

We now proceed to the study of the performance of the exact maximum likelihood method in the presence of an $AR(1)$ component. To do this, we simulate $ARFIMA(1, d, 0)$ processes of length 300. The values for the autoregressive parameter are spread between -0.9 and 0.9, with a step of 0.2.

3.2.1 General comments

Tables 5 and 6 report the number of times the right model has been selected by AIC and SIC criteria respectively.

Results in tables 5 and 6 indicate that both information criteria perform well, since $ARFIMA(1, d, 0)$ is the most often selected model, except for $\phi = 0.1$ and -0.1. The SIC criterion often selects

Table 7: Processes most often selected by the AIC criterion

						d					
ϕ	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5
-0.9	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,1)	(1,d,1)
-0.7	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)
-0.5	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)
-0.3	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)
-0.1	(2,d,2)	(2,d,2)	(0,d,0)	(1,d,0)	(1,d,0)	(0,d,0)	(0,d,0)	(0,d,0)	(3,d,2)	(1,d,0)	(1,d,0)
0.1	(0,d,0)	(2,d,2)	(2,d,2)	(2,d,2)	(2,d,2)	(2,d,2)	(0,d,0)	(0,d,0)	(0,d,0)	(0,d,0)	(0,d,0)
0.3	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)
0.5	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)
0.7	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)
0.9	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(2,d,0)

Table 8: Processes most often selected by the SIC criterion

						d					
ϕ	-0.5	-0.4	-0.3	-0.2	-0.1	0	0.1	0.2	0.3	0.4	0.5
-0.9	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,1)
-0.7	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)
-0.5	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)
-0.3	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)
-0.1	(1,d,0)	(0,d,0)	(0,d,0)	(1,d,0)	(0,d,0)	(0,d,0)	(0,d,0)	(0,d,0)	(0,d,0)	(0,d,0)	(1,d,0)
0.1	(0,d,0)	(0,d,0)	(1,d,0)	(0,d,0)	(0,d,0)	(0,d,0)	(0,d,0)	(0,d,0)	(0,d,0)	(0,d,0)	(0,d,0)
0.3	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)
0.5	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)
0.7	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)
0.9	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)	(1,d,0)

the right model, especially for high ϕ values. Globally, according to tables 5 and 6, SIC performs better than AIC. This conclusion contrasts with the results obtained by Crato and Ray (1996) who found that AIC produces better results for $ARFIMA(1,d,0)$ processes with Geweke and Porter-Hudak and approximate maximum likelihood estimation procedures. However, as noted by Crato and Ray (1996), even the AIC criterion had a very low success rate. This illustrates that selection criteria may behave differently with different estimation procedures when dealing with ARFIMA processes which also incorporate short-range dependence.

Moreover, compared to Crato and Ray (1996)'s results, our number of successful identifications is very high. In their study, this number was extremely low. As they selected short memory ARMA models, it indicates that short memory component and long-range dependence cannot be distinguished in small and medium samples. Once again, this conclusion shows that success in model selection criteria depends upon the estimation procedure.

Results in table 7 indicate that when $\phi = 0.1$ and for positive d values, AIC is biased towards a pure fractionally integrated process, *i.e.* a process without short-range dependence. For the same ϕ value but for negative d values, except for $d = -0.5$, AIC tends to select an $ARFIMA(2,d,2)$ process, *i.e.* it overparametrizes ARFIMA models. In other cases, AIC generally selects the right model. Whatever the value of d (except for $d = -0.3$), the SIC criterion is biased towards $ARFIMA(0,d,0)$ for $|\phi| = 0.1$ (see table 8). For the other values, SIC performs well and selects the right model.

3.2.2 Bias and RMSE associated with the fractional differencing parameter

We can now proceed to a detailed study of bias and RMSE concerning the fractional differencing parameter. In order to avoid an overabundance of tables, we only report some results for particular values of ϕ and d (see appendix, tables 9 to 20)[8]. As previously, we consider three cases: Bias and RMSE associated with the model selected by the considered criterion (case (1)), bias and RMSE corresponding to the right model which has been selected (case (2)), and bias and RMSE associated with the right model (even if this model had not been selected by the considered criterion, case (3)).

Concerning bias, various comments can be made (see tables 9 to 12). Generally, and as for pure fractional noise models, the fractional differencing parameter is underestimated since the bias is negative. One main exception concerns $\phi = -0.9$. Indeed, if one considers case (2), AIC produces a positive bias for negative d values, and so does SIC except for strong persistent processes ($d = 0.5$ and 0.4).

The analysis of results for case (1) shows that the minimal bias is reached by SIC except for $|\phi| = 0.9$ and strong anti-persistent processes. Indeed, in this last case, the minimal bias is obtained with the AIC criterion. Concerning now case (2), results are less clear-cut, and depend upon the signs of both ϕ and d. More specifically, results indicate that:

- For high positive ϕ values, AIC frequently leads to the lowest bias, except for strong persistent and anti-persistent processes (*i.e.* for high $|d|$ values). When ϕ decreases, minimal bias is generally obtained with SIC. Finally, for low ϕ values (*i.e.* for $\phi = 0.1$), the minimal bias is globally reached with SIC for negative d values, and with AIC for positive d values.

- Reciprocally, for high negative ϕ values ($\phi = -0.7$), SIC often leads to the lowest bias. For less negative ϕ values, minimal bias tends to be reached by AIC. Finally, when $\phi = -0.1$, the lowest bias is generally associated with SIC for negative d values, and with AIC for positive d values.

Let us now proceed to the analysis of the relation between bias and d according to each criterion. Two general comments can be made:

- For the AIC criterion, bias also increases as d increases, especially in case (1). If one considers case (2), bias seems to be independent of d for high and low $|\phi|$ values. For intermediate ϕ values, bias slightly increases along with the value of d.

- For the SIC criterion, and whatever the case, bias does not depend on d for $|\phi| = 0.9$. This is also the case for low $|\phi|$ values. For intermediate ϕ values, bias tends to increase with d.

To conclude, one can remark that, whatever the criterion, the bias is generally lower for positive ϕ values than for negative ones. In other words, the underestimation of d is more important in the presence of a negative short-term parameter.

Consider now tables 13 to 16 relating to RMSE results. Results relating to case (1) show that SIC generally leads to the lowest RMSE, except in the case of strong anti-persistent processes when $\phi = -0.9$ and where minimal RMSE is reached by AIC. If we now consider the case when the right model has been selected by information criteria (case (2)), results are more disparate and depend upon the signs and values of d and ϕ. More specifically, for positive ϕ values, one has the following pattern:

- For high values of ϕ, the lowest RMSE is given by AIC for positive d values, and by SIC for negative ones.

- For decreasing ϕ values, AIC generally leads to the lowest RMSE.

- When ϕ reaches 0.3, SIC outperforms AIC.

[8] The other tables are available upon request to authors.

- Finally, for $\phi = 0.1$, AIC leads to better results for positive d values, and SIC outperforms AIC for negative d values.

For negative ϕ values, the following results hold:

- For low ϕ in absolute value, RMSE is generally minimized by AIC, except for strong persistent processes ($d = 0.5$) and for some negative d values.
- When ϕ increases (in absolute value), SIC leads to the lowest RMSE.
- Finally, for $\phi = -0.9$, AIC is associated with the lowest RMSE, except for high positive d values.

Some comments on the relationship between RMSE and d must be added. In case (1), RMSE tends to increase with d for both information criteria. When the right model is selected (case (2)), RMSE is not very sensitive to d, except in the following two cases: for low and positive ϕ values, RMSE tends to increase as d increases, and, reciprocally, for high and negative ϕ values, RMSE tends to decrease as d increases.

Finally, if one studies RMSE values for various ϕ values, it appears that the lowest RMSE is always found when combined with positive ϕ values. This result confirms the conclusion drawn from bias analysis.

3.2.3 Bias and RMSE associated with the short-term component

The analysis of bias and RMSE associated with the autoregressive coefficient leads to the following results (see tables 17 to 20). First, bias is generally very low. Moreover, it is frequently negative, except for $\phi = -0.9$. Second, in case (1), SIC gives the lowest bias. If we consider the case when the right model is selected (case (2)), results depend upon the signs of d and ϕ. More specifically, the following pattern can be observed:

- When $\phi < 0$ and $d < 0$, SIC generally outperforms AIC.
- When $\phi > 0$, AIC frequently leads to the lowest bias.
- For low positive d values, AIC tends to outperform SIC in terms of bias.
- Finally, for high d values, SIC tends to outperform AIC.

The third and last conclusion concerns RMSE of ϕ. RMSE is always minimized by the SIC criterion if one only considers case (1). In case (2), SIC outperforms AIC for high negative d values. For low positive d values, AIC tends to minimize RMSE more frequently than SIC. Finally, for high positive d values, AIC leads to the lowest RMSE when $\phi > 0$, while SIC minimizes RMSE for $\phi < 0$.

4 Conclusion

We have resorted to a simulation study to assess the performance of two automatic selection criteria when long-term memory processes are estimated by the exact maximum likelihood method. Three main conclusions emerge from this Monte Carlo analysis. First, whatever the criterion, the bias is generally negative, which means that d tends to be underestimated. This is a well known result on EML. Moreover, results relating to $ARFIMA(1, d, 0)$ processes show that bias and RMSE are generally lower with a positive first-order autoregressive coefficient than with a negative one. However, contrary to Hauser (1999), we do not find strong evidence that bias tends to increase with the introduction of an $AR(1)$ component. Second, our results indicate that SIC outperforms AIC, in the cases of both pure fractional noise and ARFIMA processes. This criterion generally had a high success rate to select the right model. Third, if we compare our study to other works, our results show that success in the use of selection criteria is dependent on the estimation procedure when mixed ARFIMA processes are estimated. Crato and Ray (1996) showed that AIC provides the best results when Haslett and Raftery (1989) and Fox and Taqqu (1986)'s methods are used. They showed that the Geweke and Porter-Hudak procedure presents significant estimation biases and their simulation results confirm a poor model identification. Our results indicate that SIC always leads to better results for both fractional noise and ARFIMA processes when the EML method is used.

Appendix: some results of Monte Carlo simulations

Tables 9 to 12 and tables 13 to 16 respectively give bias and RMSE for the fractional differencing parameter d. Tables 17 and 18 concern bias of the first-order autoregressive coefficient. Finally, tables 19 and 20 report RMSE of ϕ.

For all tables, the legend reads as follows. (1) means bias (or RMSE) given by the selected model. (2) means bias (or RMSE) given by the selected right model. (3) means bias (or RMSE) for the right model. In italics: minimum bias (or RMSE) in (1). In bold: minimum bias (or RMSE) in (2).

Table 9: ARFIMA(1,d,0) — Estimation of d with $\phi = -0.7$ — Bias

	AIC		SIC		
d	(1)	(2)	(1)	(2)	(3)
-0.5	$-1.01.10^{-1}$	$-2.87.10^{-2}$	*$-1.1.10^{-2}$*	**$-2.18.10^{-2}$**	$-2.42.10^{-3}$
-0.4	$-1.36.10^{-1}$	$-3.39.10^{-2}$	*$-2.65.10^{-2}$*	**$-2.8.10^{-2}$**	$-1.07.10^{-2}$
-0.3	$-1.72.10^{-1}$	$-4.22.10^{-2}$	*$-4.74.10^{-2}$*	**$-3.65.10^{-2}$**	$-2.06.10^{-2}$
-0.2	$-2.49.10^{-1}$	$-5.8.10^{-2}$	*$-6.91.10^{-2}$*	**$-5.55.10^{-2}$**	$-4.18.10^{-2}$
-0.1	$-2.73.10^{-1}$	$-7.47.10^{-2}$	*$-9.97.10^{-2}$*	**$-6.85.10^{-2}$**	$-5.9.10^{-2}$
0.0	$-2.99.10^{-1}$	$-9.78.10^{-2}$	*$-1.13.10^{-1}$*	**$-8.88.10^{-2}$**	$-8.4.10^{-2}$
0.1	$-3.35.10^{-1}$	$-9.22.10^{-2}$	*$-1.29.10^{-1}$*	**$-8.29.10^{-2}$**	$-1.07.10^{-1}$
0.2	$-4.01.10^{-1}$	$-1.06.10^{-1}$	*$-1.5.10^{-1}$*	**$-9.82.10^{-2}$**	$-1.9.10^{-1}$
0.3	$-4.34.10^{-1}$	$-1.17.10^{-1}$	*$-1.7.10^{-1}$*	**$-1.02.10^{-1}$**	$-1.75.10^{-1}$
0.4	$-4.61.10^{-1}$	$-1.13.10^{-1}$	*$-1.76.10^{-1}$*	**$-1.09.10^{-1}$**	$-1.25.10^{-1}$
0.5	$-4.51.10^{-1}$	$-1.57.10^{-1}$	*$-2.1.10^{-1}$*	**$-1.44.10^{-1}$**	$-1.72.10^{-1}$

Table 10: ARFIMA(1,d,0) — Estimation of d with $\phi = -0.1$ — Bias

	AIC		SIC		
d	(1)	(2)	(1)	(2)	(3)
-0.5	-7.16.10^{-2}	-1.12.10^{-1}	*-1.63.10^{-2}*	**-6.65.10^{-2}**	-4.71.10^{-2}
-0.4	-1.03.10^{-1}	-1.41.10^{-1}	*-3.36.10^{-3}*	**-1.07.10^{-1}**	-5.18.10^{-2}
-0.3	-1.23.10^{-1}	-1.52.10^{-1}	*-1.04.10^{-2}*	**-1.32.10^{-1}**	-5.99.10^{-2}
-0.2	-1.41.10^{-1}	-1.18.10^{-1}	*-4.03.10^{-2}*	**-1.06.10^{-1}**	-6.6.10^{-2}
-0.1	-1.57.10^{-1}	**-1.53.10^{-1}**	*3.82.10^{-3}*	-2.55.10^{-1}	-7.14.10^{-2}
0.0	-1.82.10^{-1}	**-2.39.10^{-1}**	*-1.53.10^{-2}*	-3.76.10^{-1}	-8.29.10^{-2}
0.1	-2.13.10^{-1}	**-2.42.10^{-1}**	*-2.22.10^{-2}*	-3.89.10^{-1}	-8.62.10^{-2}
0.2	-2.55.10^{-1}	**-2.59.10^{-1}**	*-3.08.10^{-2}*	-3.95.10^{-1}	-9.5.10^{-2}
0.3	-3.09.10^{-1}	**-3.08.10^{-1}**	*-4.88.10^{-2}*	-4.43.10^{-1}	-1.12.10^{-1}
0.4	-3.87.10^{-1}	**-3.56.10^{-1}**	*-1.06.10^{-1}*	-4.56.10^{-1}	-1.56.10^{-1}
0.5	-6.42.10^{-1}	-6.62.10^{-1}	*-3.83.10^{-1}*	**-5.74.10^{-1}**	-4.51.10^{-1}

Table 11: ARFIMA(1,d,0) — Estimation of d with $\phi = 0.1$ — Bias

	AIC		SIC		
d	(1)	(2)	(1)	(2)	(3)
-0.5	-6.54.10^{-2}	**-1.83.10^{-2}**	*-4.59.10^{-2}*	2.35.10^{-2}	-1.85.10^{-2}
-0.4	-9.96.10^{-2}	-2.09.10^{-2}	*-4.66.10^{-2}*	**-7.29.10^{-3}**	-2.71.10^{-2}
-0.3	-1.15.10^{-1}	-1.85.10^{-2}	*-3.89.10^{-2}*	**-1.73.10^{-2}**	-3.09.10^{-2}
-0.2	-1.46.10^{-1}	-2.31.10^{-2}	*-4.71.10^{-2}*	**-2.31.10^{-2}**	-3.4.10^{-2}
-0.1	-1.54.10^{-1}	**-1.41.10^{-2}**	*-6.86.10^{-2}*	-1.62.10^{-2}	-3.84.10^{-2}
0.0	-1.73.10^{-1}	**-3.16.10^{-2}**	*-4.59.10^{-2}*	-4.94.10^{-2}	-3.93.10^{-2}
0.1	-2.05.10^{-1}	-2.62.10^{-2}	*-6.63.10^{-2}*	**-8.72.10^{-3}**	-4.19.10^{-2}
0.2	-2.42.10^{-1}	**-1.52.10^{-2}**	*-7.92.10^{-2}*	1.64.10^{-2}	-4.51.10^{-2}
0.3	-2.79.10^{-1}	**-1.75.10^{-2}**	*-8.57.10^{-2}*	2.38.10^{-2}	-4.96.10^{-2}
0.4	-3.49.10^{-1}	**-6.26.10^{-2}**	*-1.01.10^{-1}*	-1.02.10^{-1}	-6.42.10^{-2}
0.5	-4.81.10^{-1}	**-3.56.10^{-1}**	*-1.34.10^{-1}*	-4.57.10^{-1}	-9.88.10^{-2}

Table 12: ARFIMA(1,d,0) — Estimation of d with $\phi = 0.7$ — Bias

	AIC		SIC		
d	(1)	(2)	(1)	(2)	(3)
-0.5	-1.4.10^{-2}	**-6.97.10^{-4}**	*1.7.10^{-3}*	1.58.10^{-3}	1.35.10^{-3}
-0.4	-4.79.10^{-2}	-9.13.10^{-3}	*-1.01.10^{-2}*	**-8.76.10^{-3}**	-8.88.10^{-3}
-0.3	-8.8.10^{-2}	-1.52.10^{-2}	*-1.99.10^{-2}*	**-1.38.10^{-2}**	-1.38.10^{-2}
-0.2	-9.94.10^{-2}	-1.72.10^{-2}	*-2.65.10^{-2}*	**-1.67.10^{-2}**	-1.69.10^{-2}
-0.1	-1.34.10^{-1}	**-1.98.10^{-2}**	*-3.19.10^{-2}*	-2.01.10^{-2}	-2.10^{-2}
0.0	-1.48.10^{-1}	**-2.05.10^{-2}**	*-3.98.10^{-2}*	-2.1.10^{-2}	-2.02.10^{-2}
0.1	-1.62.10^{-1}	**-1.87.10^{-2}**	*-4.21.10^{-2}*	-2.12.10^{-2}	-2.1.10^{-2}
0.2	-2.23.10^{-1}	**-2.24.10^{-2}**	*-5.8.10^{-2}*	-2.46.10^{-2}	-2.41.10^{-2}
0.3	-2.67.10^{-1}	**-2.58.10^{-2}**	*-6.57.10^{-2}*	-2.74.10^{-2}	-2.72.10^{-2}
0.4	-3.41.10^{-1}	**-3.14.10^{-2}**	*-9.13.10^{-2}*	-3.32.10^{-2}	-3.25.10^{-2}
0.5	-4.35.10^{-1}	-6.02.10^{-2}	*-1.41.10^{-1}*	**-5.5.10^{-2}**	-5.41.10^{-2}

Table 13: ARFIMA(1,d,0) — Estimation of d with $\phi = -0.7$ — RMSE

	AIC		SIC		
d	(1)	(2)	(1)	(2)	(3)
-0.5	$3.22.10^{-1}$	$1.39.10^{-1}$	*$2.04.10^{-1}$*	**$1.36.10^{-1}$**	$1.5.10^{-1}$
-0.4	$3.47.10^{-1}$	$1.45.10^{-1}$	*$2.08.10^{-1}$*	**$1.36.10^{-1}$**	$1.46.10^{-1}$
-0.3	$3.71.10^{-1}$	$1.46.10^{-1}$	*$2.23.10^{-1}$*	**$1.35.10^{-1}$**	$1.43.10^{-1}$
-0.2	$4.62.10^{-1}$	$1.48.10^{-1}$	*$2.29.10^{-1}$*	**$1.4.10^{-1}$**	$1.44.10^{-1}$
-0.1	$4.72.10^{-1}$	$1.5.10^{-1}$	*$2.48.10^{-1}$*	**$1.42.10^{-1}$**	$1.43.10^{-1}$
0.0	$4.89.10^{-1}$	$1.61.10^{-1}$	*$2.31.10^{-1}$*	**$1.54.10^{-1}$**	$1.55.10^{-1}$
0.1	$5.24.10^{-1}$	$1.47.10^{-1}$	*$2.78.10^{-1}$*	**$1.41.10^{-1}$**	$1.85.10^{-1}$
0.2	$5.74.10^{-1}$	$1.53.10^{-1}$	*$2.8.10^{-1}$*	**$1.46.10^{-1}$**	$2.87.10^{-1}$
0.3	$5.94.10^{-1}$	$1.62.10^{-1}$	*$2.75.10^{-1}$*	**$1.49.10^{-1}$**	$2.76.10^{-1}$
0.4	$6.14.10^{-1}$	$1.46.10^{-1}$	*$2.73.10^{-1}$*	**$1.4.10^{-1}$**	$1.89.10^{-1}$
0.5	$5.79.10^{-1}$	$1.72.10^{-1}$	*$2.89.10^{-1}$*	**$1.59.10^{-1}$**	$2.18.10^{-1}$

Table 14: ARFIMA(1,d,0) — Estimation of d with $\phi = -0.1$ — RMSE

	AIC		SIC		
d	(1)	(2)	(1)	(2)	(3)
-0.5	$2.63.10^{-1}$	$2.56.10^{-1}$	*$1.52.10^{-1}$*	**$2.05.10^{-1}$**	$1.68.10^{-1}$
-0.4	$2.9.10^{-1}$	$2.52.10^{-1}$	*$1.59.10^{-1}$*	**$2.4.10^{-1}$**	$1.49.10^{-1}$
-0.3	$3.01.10^{-1}$	**$2.46.10^{-1}$**	*$1.48.10^{-1}$*	$2.53.10^{-1}$	$1.51.10^{-1}$
-0.2	$3.05.10^{-1}$	$2.18.10^{-1}$	*$1.82.10^{-1}$*	**$2.03.10^{-1}$**	$1.6.10^{-1}$
-0.1	$3.33.10^{-1}$	**$2.45.10^{-1}$**	*$1.81.10^{-1}$*	$3.34.10^{-1}$	$1.68.10^{-1}$
0.0	$3.7.10^{-1}$	**$3.32.10^{-1}$**	*$1.82.10^{-1}$*	$4.49.10^{-1}$	$1.87.10^{-1}$
0.1	$4.06.10^{-1}$	**$3.4.10^{-1}$**	*$1.91.10^{-1}$*	$4.66.10^{-1}$	$1.95.10^{-1}$
0.2	$4.53.10^{-1}$	**$3.58.10^{-1}$**	*$1.98.10^{-1}$*	$4.7.10^{-1}$	$2.01.10^{-1}$
0.3	$4.94.10^{-1}$	**$4.21.10^{-1}$**	*$2.26.10^{-1}$*	$5.28.10^{-1}$	$2.32.10^{-1}$
0.4	$5.64.10^{-1}$	**$4.86.10^{-1}$**	*$2.85.10^{-1}$*	$5.64.10^{-1}$	$2.93.10^{-1}$
0.5	$7.48.10^{-1}$	$7.4.10^{-1}$	*$5.43.10^{-1}$*	**$6.84.10^{-1}$**	$5.84.10^{-1}$

Table 15: ARFIMA(1,d,0) — Estimation of d with $\phi = 0.1$ — RMSE

	AIC		SIC		
d	(1)	(2)	(1)	(2)	(3)
-0.5	$2.5.10^{-1}$	$2.24.10^{-1}$	*$1.02.10^{-1}$*	**$1.23.10^{-1}$**	$9.99.10^{-2}$
-0.4	$2.65.10^{-1}$	$1.65.10^{-1}$	*$1.19.10^{-1}$*	**$1.08.10^{-1}$**	$9.82.10^{-2}$
-0.3	$2.8.10^{-1}$	**$8.19.10^{-2}$**	*$1.27.10^{-1}$*	$8.95.10^{-2}$	$8.95.10^{-2}$
-0.2	$3.06.10^{-1}$	$9.84.10^{-2}$	*$1.15.10^{-1}$*	**$8.84.10^{-2}$**	$8.7.10^{-2}$
-0.1	$3.13.10^{-1}$	$1.1.10^{-1}$	*$1.44.10^{-1}$*	**$1.05.10^{-1}$**	$9.45.10^{-2}$
0.0	$3.44.10^{-1}$	**$1.13.10^{-1}$**	*$1.36.10^{-1}$*	$1.36.10^{-1}$	$9.95.10^{-2}$
0.1	$3.75.10^{-1}$	**$1.5.10^{-1}$**	*$1.2.10^{-1}$*	$2.14.10^{-1}$	$1.01.10^{-1}$
0.2	$4.16.10^{-1}$	**$1.62.10^{-1}$**	*$1.33.10^{-1}$*	$2.11.10^{-1}$	$1.01.10^{-1}$
0.3	$4.56.10^{-1}$	**$1.51.10^{-1}$**	*$1.42.10^{-1}$*	$1.63.10^{-1}$	$9.76.10^{-2}$
0.4	$5.26.10^{-1}$	**$2.07.10^{-1}$**	*$1.69.10^{-1}$*	$3.42.10^{-1}$	$1.16.10^{-1}$
0.5	$6.32.10^{-1}$	**$5.05.10^{-1}$**	*$2.18.10^{-1}$*	$6.36.10^{-1}$	$1.5.10^{-1}$

Table 16: ARFIMA(1,d,0) — Estimation of d with $\phi = 0.7$ — RMSE

	AIC		SIC		
d	(1)	(2)	(1)	(2)	(3)
-0.5	$2.22.10^{-1}$	$\mathbf{5.14.10^{-2}}$	*$7.09.10^{-2}$*	$5.17.10^{-2}$	$5.18.10^{-2}$
-0.4	$2.16.10^{-1}$	$\mathbf{4.93.10^{-2}}$	*$7.05.10^{-2}$*	$5.22.10^{-2}$	$5.22.10^{-2}$
-0.3	$2.57.10^{-1}$	$5.35.10^{-2}$	*$9.09.10^{-2}$*	$\mathbf{5.33.10^{-2}}$	$5.35.10^{-2}$
-0.2	$2.75.10^{-1}$	$\mathbf{5.29.10^{-2}}$	*$1.04.10^{-1}$*	$5.38.10^{-2}$	$5.4.10^{-2}$
-0.1	$3.15.10^{-1}$	$\mathbf{5.69.10^{-2}}$	*$1.13.10^{-1}$*	$5.71.10^{-2}$	$5.69.10^{-2}$
0.0	$3.26.10^{-1}$	$\mathbf{5.67.10^{-2}}$	*$1.36.10^{-1}$*	$5.74.10^{-2}$	$5.71.10^{-2}$
0.1	$3.49.10^{-1}$	$\mathbf{5.37.10^{-2}}$	*$1.4.10^{-1}$*	$5.59.10^{-2}$	$5.6.10^{-2}$
0.2	$4.13.10^{-1}$	$\mathbf{5.52.10^{-2}}$	*$1.76.10^{-1}$*	$5.87.10^{-2}$	$5.86.10^{-2}$
0.3	$4.52.10^{-1}$	$\mathbf{5.59.10^{-2}}$	*$1.87.10^{-1}$*	$5.89.10^{-2}$	$5.91.10^{-2}$
0.4	$5.27.10^{-1}$	$\mathbf{5.56.10^{-2}}$	*$2.28.10^{-1}$*	$5.88.10^{-2}$	$5.86.10^{-2}$
0.5	$5.85.10^{-1}$	$6.79.10^{-2}$	*$2.84.10^{-1}$*	$\mathbf{6.46.10^{-2}}$	$6.4.10^{-2}$

Table 17: ARFIMA(1,d,0) — Estimation of ϕ with $d = -0.4$ — Bias

	AIC		SIC		
ϕ	(1)	(2)	(1)	(2)	(3)
-0.9	*$2.32.10^{-2}$*	$6.58.10^{-2}$	$1.98.10^{-1}$	$\mathbf{6.48.10^{-2}}$	$2.05.10^{-1}$
-0.7	$-2.71.10^{-2}$	$-4.84.10^{-3}$	*$1.47.10^{-2}$*	$\mathbf{-3.17.10^{-3}}$	$1.36.10^{-2}$
-0.5	$3.47.10^{-2}$	$-1.12.10^{-1}$	*$-5.71.10^{-3}$*	$\mathbf{-9.89.10^{-2}}$	$-7.19.10^{-2}$
-0.3	*$-1.82.10^{-2}$*	$-1.38.10^{-1}$	$-1.93.10^{-2}$	$\mathbf{-1.35.10^{-1}}$	$-8.81.10^{-2}$
-0.1	*$-1.92.10^{-2}$*	$-1.4.10^{-1}$	$4.89.10^{-2}$	$\mathbf{-8.52.10^{-2}}$	$-4.67.10^{-2}$
0.1	$-1.01.10^{-1}$	$-1.95.10^{-2}$	*$-6.63.10^{-2}$*	$\mathbf{-1.56.10^{-2}}$	$-2.12.10^{-2}$
0.3	$-1.19.10^{-1}$	$\mathbf{7.2.10^{-4}}$	*$-7.5.10^{-2}$*	$1.53.10^{-2}$	$-1.12.10^{-2}$
0.5	$-9.5.10^{-2}$	$\mathbf{-7.98.10^{-3}}$	*$-3.69.10^{-2}$*	$-1.1.10^{-2}$	$-1.17.10^{-2}$
0.7	$-6.94.10^{-2}$	$-1.35.10^{-2}$	*$-1.12.10^{-2}$*	$\mathbf{-1.32.10^{-2}}$	$-1.33.10^{-2}$
0.9	*$-4.41.10^{-2}$*	$-5.01.10^{-2}$	$-4.54.10^{-2}$	$\mathbf{-4.79.10^{-2}}$	$-4.69.10^{-2}$

Table 18: ARFIMA(1,d,0) — Estimation of ϕ with $d = 0.4$ — Bias

	AIC		SIC		
ϕ	(1)	(2)	(1)	(2)	(3)
-0.9	$-2.32.10^{-1}$	$\mathbf{7.25.10^{-3}}$	*$-2.41.10^{-2}$*	$1.95.10^{-2}$	$-1.92.10^{-2}$
-0.7	$-2.83.10^{-1}$	$-6.77.10^{-2}$	*$-1.08.10^{-1}$*	$\mathbf{-6.28.10^{-2}}$	$-7.14.10^{-2}$
-0.5	$-3.41.10^{-1}$	$-2.04.10^{-1}$	*$-1.64.10^{-1}$*	$\mathbf{-1.96.10^{-1}}$	$-2.77.10^{-1}$
-0.3	$-3.9.10^{-1}$	$-4.19.10^{-1}$	*$-2.35.10^{-1}$*	$\mathbf{-4.14.10^{-1}}$	$-4.52.10^{-1}$
-0.1	$-2.54.10^{-1}$	$\mathbf{-3.69.10^{-1}}$	*$-4.33.10^{-2}$*	$-4.83.10^{-1}$	$-1.47.10^{-1}$
0.1	$-3.08.10^{-1}$	$\mathbf{-2.9.10^{-2}}$	*$-1.09.10^{-1}$*	$-3.41.10^{-2}$	$-5.58.10^{-2}$
0.3	$-3.22.10^{-1}$	$-2.5.10^{-2}$	*$-1.31.10^{-1}$*	$\mathbf{3.99.10^{-5}}$	$-3.14.10^{-2}$
0.5	$-2.59.10^{-1}$	$\mathbf{-2.04.10^{-2}}$	*$-8.59.10^{-2}$*	$-2.38.10^{-2}$	$-2.45.10^{-2}$
0.7	$-2.57.10^{-1}$	$-1.75.10^{-2}$	*$-7.52.10^{-2}$*	$\mathbf{-1.72.10^{-2}}$	$-1.69.10^{-2}$
0.9	$-3.48.10^{-1}$	$\mathbf{-1.31.10^{-2}}$	*$-7.45.10^{-2}$*	$-1.56.10^{-2}$	$-1.64.10^{-2}$

Table 19: ARFIMA(1,d,0) — Estimation of ϕ with $d = -0.4$ — RMSE

	AIC		SIC		
ϕ	(1)	(2)	(1)	(2)	(3)
-0.9	$6.43.10^{-1}$	$1.45.10^{-1}$	$\mathit{3.97.10^{-1}}$	$\mathbf{1.4.10^{-1}}$	$3.64.10^{-1}$
-0.7	$6.72.10^{-1}$	$1.26.10^{-1}$	$\mathit{2.38.10^{-1}}$	$\mathbf{1.14.10^{-1}}$	$1.34.10^{-1}$
-0.5	$7.11.10^{-1}$	$1.94.10^{-1}$	$\mathit{2.74.10^{-1}}$	$\mathbf{1.78.10^{-1}}$	$1.73.10^{-1}$
-0.3	$7.5.10^{-1}$	$2.48.10^{-1}$	$\mathit{2.62.10^{-1}}$	$\mathbf{2.41.10^{-1}}$	$2.16.10^{-1}$
-0.1	$7.81.10^{-1}$	$2.57.10^{-1}$	$\mathit{1.83.10^{-1}}$	$\mathbf{2.53.10^{-1}}$	$1.6.10^{-1}$
0.1	$7.96.10^{-1}$	$1.75.10^{-1}$	$\mathit{1.42.10^{-1}}$	$\mathbf{1.31.10^{-1}}$	$1.12.10^{-1}$
0.3	$7.09.10^{-1}$	$6.63.10^{-2}$	$\mathit{1.92.10^{-1}}$	$\mathbf{5.93.10^{-2}}$	$7.59.10^{-2}$
0.5	$6.16.10^{-1}$	$\mathbf{6.31.10^{-2}}$	$\mathit{1.44.10^{-1}}$	$6.34.10^{-2}$	$6.39.10^{-2}$
0.7	$5.81.10^{-1}$	$4.93.10^{-2}$	$\mathit{8.72.10^{-2}}$	$\mathbf{4.82.10^{-2}}$	$4.84.10^{-2}$
0.9	$3.89.10^{-1}$	$5.44.10^{-2}$	$\mathit{7.10^{-2}}$	$\mathbf{5.29.10^{-2}}$	$5.23.10^{-2}$

Table 20: ARFIMA(1,d,0) — Estimation of ϕ with $d = 0.4$ — RMSE

	AIC		SIC		
ϕ	(1)	(2)	(1)	(2)	(3)
-0.9	$6.21.10^{-1}$	$\mathbf{3.38.10^{-2}}$	$\mathit{1.94.10^{-1}}$	$4.14.10^{-2}$	$6.34.10^{-2}$
-0.7	$6.4.10^{-1}$	$9.83.10^{-2}$	$\mathit{2.19.10^{-1}}$	$\mathbf{9.11.10^{-2}}$	$1.08.10^{-1}$
-0.5	$6.92.10^{-1}$	$2.43.10^{-1}$	$\mathit{3.58.10^{-1}}$	$\mathbf{2.37.10^{-1}}$	$3.28.10^{-1}$
-0.3	$7.86.10^{-1}$	$4.76.10^{-1}$	$\mathit{4.61.10^{-1}}$	$\mathbf{4.72.10^{-1}}$	$5.18.10^{-1}$
-0.1	$8.25.10^{-1}$	$\mathbf{4.94.10^{-1}}$	$\mathit{3.04.10^{-1}}$	$5.77.10^{-1}$	3.10^{-1}
0.1	$8.61.10^{-1}$	$\mathbf{2.34.10^{-1}}$	$\mathit{1.68.10^{-1}}$	$3.81.10^{-1}$	$1.3.10^{-1}$
0.3	$8.35.10^{-1}$	7.10^{-2}	$\mathit{2.85.10^{-1}}$	$\mathbf{5.38.10^{-2}}$	$8.28.10^{-2}$
0.5	$7.04.10^{-1}$	$\mathbf{6.46.10^{-2}}$	$\mathit{2.37.10^{-1}}$	$6.79.10^{-2}$	$6.77.10^{-2}$
0.7	$6.79.10^{-1}$	$\mathbf{5.1.10^{-2}}$	$\mathit{2.32.10^{-1}}$	$5.14.10^{-2}$	$5.1.10^{-2}$
0.9	$6.63.10^{-1}$	$\mathbf{3.11.10^{-2}}$	$\mathit{2.43.10^{-1}}$	$3.23.10^{-2}$	$3.29.10^{-2}$

References

[1] Agiakloglou C., Newbold P. and Wohar M. (1993), "Bias in an Estimator of the Fractional Difference Parameter", *Journal of Time Series Analysis* 14, 235-246.

[2] An S., Bloomfield P. and Pantula S. (1992), "Asymptotic Properties of the MLE in Fractional ARIMA models", Institute of Statistics Mimeograph Series No. 2228, North California State University.

[3] Andrews D.W.K. and Guggenberger P. (2003), "A Bias-reduced Log-Periodogram Regression Estimator for the Long-Memory Parameter", *Econometrica* 71, 675-712.

[4] Andrews D.W.K. and Sun Y. (2004), "Additive Local Polynomial Whittle Estimation of Long-Range Dependence", *Econometrica* 72, 569-614.

[5] Baillie R.T. (1996), "Long Memory Processes and Fractional Integration in Econometrics", *Journal of Econometrics* 73(1), 5-59.

[6] Beran J., Bhansali R.J. and Ocker D. (1998), "On Unified Model Selection for Stationary and Nonstationary Short- and Long-Memory Autoregressive Processes", *Biometrika* 85, 921-934.

[7] Bollerslev T. and Jubinski D. (1999), "Equity Trading Volume and Volatility: Latent Information Arrivals and Common Long-run Dependencies", *Journal of Business and Economic Statistics* 17, 9-21.

[8] Cheung Y.-W. and Diebold F.X. (1994), "On the Maximum Likelihood Estimation of the Differencing Parameter of Fractionally Integrated Noise with Unknown Mean", *Journal of Econometrics* 62, 301-316.

[9] Crato N. and Ray B.K. (1996), "Model Selection and Forecasting for Long-range Dependent Processes", *Journal of Forecasting* 15, 107-125.

[10] Dahlhaus R. (1988), "Small Sample Effects in Time Series Analysis: A New Asymptotic Theory and a New Estimate", *The Annals of Statistics* 16, 808-841.

[11] Dahlhaus R. (1989), "Efficient Parameter Estimation for Self-Similar Processes", *The Annals of Statistics* 17, 1749-1766.

[12] Doornik J.A. and Ooms M. (2003), "Computational Aspects of Maximum Likelihood Estimation of Autoregressive Fractionally Integrated Moving Average Models", *Computational Statistics and Data Analysis* 41, 333-348.

[13] Fox R. and Taqqu M.S. (1986), "Large-Sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series", *Annals of Statistics* 14, 517-532.

[14] Geweke J. and Porter-Hudak S. (1983), "The Estimation and Application of Long Memory Time Series Models", *Journal of Time Series Analysis* 4(4), 221-238.

[15] Granger C.W.J. and Joyeux R. (1980), "An Introduction to Long-Memory Time Series Models and Fractional Differencing", *Journal of Time Series Analysis* 1(1), 15-29.

[16] Haslett J. and Raftery A.E. (1989), "Space Time Modelling with Long Memory Dependence: Assessing Ireland's Wind Power Resource", *Applied Statistics* 38, 1-50.

[17] Hauser M.A. (1999), "Maximum Likelihood Estimators for ARMA and ARFIMA Models: A Monte Carlo Study", *Journal of Statistical Planning and Inference* 80, 229-255.

[18] Hosking J.R.M. (1981), "Fractional Differencing", *Biometrika* 68(1), 165-176.

[19] Janacek G.J. (1982), "Determining the Degree for Time Series via the Log Spectrum", *Journal of Time Series Analysis* 3, 177-183.

[20] Künsch H.R. (1987), “Statistical Aspects of Self-Similar Processes”, in Y. Prokhorov and V.V. Sazanov (eds), *Proceedings of the First World Congress of the Bernoulli Society*, VNU Science Press, Utrecht, 67-74.

[21] Lardic S. and Mignon V. (1999), “La mémoire longue en économie : une revue de la littérature”, *Journal de la Société Française de Statistique* 140 (2), 5-48.

[22] Li W.K. and McLeod A.I. (1986), “Fractional Time Series Modelling”, *Biometrika* 73(1), 217-221.

[23] Robinson P.M. (1995), “Gaussian Semiparametric Estimation of Long Range Dependence”, *Annals of Statistics* 23, 1630-1661.

[24] Robinson P.M. (2003), “Long-Memory Time Series”, in P.M. Robinson (ed), *Time Series with Long Memory*, Oxford University Press, Oxford, 4-32.

[25] Schmidt C.M. and Tschernig R. (1993), “Identification of Fractional ARIMA Models in the Presence of Long Memory”, Discussion Paper, University of Munich.

[26] Shea G.S. (1991), “Uncertainty and Implied Variance Bounds in Long-Memory Models of the Interest Rate Term Structure”, *Empirical Economics* 16, 287-312.

[27] Sowell F. (1992), “Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models”, *Journal of Econometrics* 53, 165-188.