A SETAR model with long-memory dynamics*

Gilles DUFRENOT[†] **Dominique GUEGAN**[‡]
Anne PEGUIN-FEISSOLLE[§]

July 21, 2003

Abstract

This paper presents a 2-regime SETAR model where the process under examination is governed by a long-memory process in the first regime and a short-memory process in the second regime. Persistence properties are studied and methods for locating the threshold parameter are proposed. Such a process presents a useful application to financial data and is applied to stock indices and individual asset prices.

Keywords: SETAR - Long-memory - ARFIMA model - Stock indices.

JEL classification: C32, C51, G12

**Corresponding author:* Anne PEGUIN-FEISSOLLE, GREQAM, Centre de la Vieille Charité, 2 rue de la Charité, 13002 Marseille, France, tel: +33.4.91.14.07.35, fax: +33.4.91.90.02.27, Email: peguin@ehess.cnrs-mrs.fr

[†] *ERUDITE, Université Paris 12 and GREQAM-CNRS, 2 rue de la Vieille Charité, 13002 Marseille, France, Email: lopaduf@aol.com*

[‡] *Ecole Normale Supérieure de Cachan, MORA-IDHE, UMR-CNRS, and Senior Academic Fellow de l'IEF, 61 avenue du Président Wilson, 94235, Cachan, Cédex, France, Email:guegan@ecogest.ens-cachan.fr*

[§] *GREQAM-CNRS, Centre de la Vieille Charité, 2 rue de la Vieille Charité, 13002 Marseille, France, Email: peguin@ehess.cnrs-mrs.fr*

1 Introduction

The last twenty years have experienced an explosion of papers reporting two key properties of the financial time series, namely the long-memory and nonlinear properties. Both are important and serve for purpose of forecastability when the markets are not efficient. The long-memory approach enables the study of the length of time it takes for a new information to be fully reflected in the prices of financial assets. Nonlinear analysis has proven to be useful in providing new statistical tools when pockets of predictability do exist. Recently, the potential of exploring these properties simultaneously has emerged in papers that try to bring the two aspects together. Empirical works have been done along two lines of research.

A first line of papers enhances a skeptic view. The long-memory property uncovered in the financial data would reflect strong biases in presence of neglected nonlinearities. Authors sharing this view argue that neglected switching dynamics or regime shifts imply spurious long-memory structures. Papers that highlight such a confusion include Hidalgo and Robinson (1996), Lobato and Savin (1997), Bos, Franses and Ooms (1999), Granger and Hyung (1999), Diebold and Inoue (2001). Diebold and Inoue (2001) provide some Monte Carlo results suggesting that spurious long-memory characterize many nonlinear models currently applied to economic data, among which the Markov switching models. A consequence of this pessimistic view is that nonlinearity and long-memory should be held jointly if one wants to be able to evaluate their relative importance.

Accordingly, a second strand of the literature has recently proposed "nonlinear long-memory" models. For instance, some authors provide a joint evidence of mean reversion over long horizons and nonlinear dynamics on exchange rate markets, by generalizing to the nonlinear framework the Beveridge Nelson decomposition (see, Clarida and Taylor (2001), Sarno and Taylor (2001)). Others propose new classes of long-memory models. For instance, Franses and Paap (2002), Franses, Van der Leij and Paap (2002) introduce CLEAR and Switching CLEAR processes, which show autocorrelation at high lags with an ACF that decays at a faster rate in the beginning in comparison to the ACF of an ARFIMA model. Guégan (2000, 2003) introduces the GIGARCH model which allies at the same time long-memory, seasonality and heteroscedasticity effects, with application to inflation rates. Fractionally integrated threshold models have also been proposed, that offers another potential application to financial data (see van Dick, Franses and Paap (2002)).

In this paper, we propose a 2-regime SETAR model with the following characteristics: in the first regime, the dynamics is that of a long-memory process (a fractional white noise) and in the second regime a white noise. One way to motivate the use of such a model from the point of view of the financial economics literature is as follows. As is known, the efficient market hypothesis implies that unexpected price changes behave as independent random draws. Doubts about the validity of the efficient market hypothesis are well documented and the arguments against the efficiency are numerous. Authors usually evoke some anomalies (such as week-end, January or firm size effects), the presence of noise traders, the presence of heterogenous agents, *etc.* Another attention-receiving factor of inefficiency is the imperfection of financial markets, with the presence of transaction costs. Typically, in presence of market frictions, the arbitrage behaviors imply some asymmetries. New information that induce mispricing errors are not arbitrated unless the deviations of the returns from their fundamental value are large enough. The presence of transaction costs implies infrequent tradings and prices are driven back to their equilibrium value more or less rapidly depending upon the magnitude of the deviations. It is today documented in the literature that in such a context mispricing errors exhibit stepwise adjustments, since they have to be large enough to offset the transaction costs. Such kind of adjustments are well described by threshold models (see Martens, Kofman and Vorst (1998), Dufrénot and Mignon (2002) and Dufrénot, Mignon and Péguin-Feissolle (2003)). This suggests a complex picture of the markets where situations of efficiency and inefficiency can alternate according to the magnitude of the prices variations. Inefficiency implies that the mispricing errors are not arbitrated. There thereby exists opportunities for positive profits, which are reflect by the presence of significant correlation in prices differentials. Here, we assume that strong correlations exist, that are a sign of a long-memory dynamics. Conversely, efficiency implies that prices differentials follow random draws (or, less strictly, a dynamics with a very short memory).

The SETAR process $(p_t)_t$ that we consider is written as follows:

$$\begin{cases} (1-B)^d \Delta p_{t-\tau} = \varepsilon_t^{(1)}, & \text{if } \Delta p_{t-l} \leq c: \quad \text{regime 1} \\ \Delta p_{t-\tau} = \varepsilon_t^{(2)}, & \text{if } \Delta p_{t-l} > c: \quad \text{regime 2,} \end{cases} \tag{1}$$

where $d \in (0, 1/2)$ is a fractional difference parameter, $\varepsilon_t^{(i)}, i = 1, 2$ are white noises with finite variances, B is the backward shift operator, l is an integer that captures the delays needed before the agents react to price changes. τ

is a lag integer. p_t is an asset or index price at time t and c is a threshold parameter. When studying the memory properties of such a process, three interesting questions are enhanced. Firstly, which memory - short or long - is captured by the autocorrelation function and the spectrum? Secondly, supposing that the threshold variable is known (here Δp_{t-l}), how can we estimate the threshold parameter c that delimitates the two regimes in terms of memory properties? Thirdly, how well is this SETAR model suited for the modeling of individual assets and stock indices?

The objective of this paper is twofold: it is firstly to investigate the statistical properties of the model (1), and secondly to provide some illustrative applications on real data.

The plan of the paper is as follows. Section 2 briefly presents the memory properties of the SETAR model including a long-memory regime. Monte Carlo simulations are provided to confirm the theoretical arguments. In section 3, we propose some methods for the location of the unknown parameter c. Section 4 contains some empirical applications to financial data. Section 5 concludes the paper.

2 The model and its memory properties

We consider a process $(X_t)_t$ that satisfies the following scheme: $\forall t$,

$$\begin{cases} (1-B)^d X_t = \varepsilon_t^{(1)}, & \text{if } X_{t-1} \leq c: \quad \text{regime 1} \\ X_t = \varepsilon_t^{(2)}, & \text{if } X_{t-1} > c: \quad \text{regime 2.} \end{cases} \tag{2}$$

We make the following assumptions :
(H_0): the process $(\varepsilon_t^{(i)})_t, i = 1, 2$ is a sequence of *i.i.d.* random $N(0, 1)$ variables.
(H_1): the long memory parameter d is such that $0 < d < 1/2$. So, in regime 1, the process is invertible and stationary.
Now, we define the following indicator function:

$$I_t(X_{t-1} \leq c) = \begin{cases} 1, & \text{if } X_{t-1} \leq c \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The $SETAR$ process $(X_t)_t$ defined in (2) can be rewritten as: $\forall t$,

$$X_t = (1-B)^{-d} \varepsilon_t^{(1)} \, I_t(X_{t-1} \leq c) + \varepsilon_t^{(2)} \left[1 - I_t(X_{t-1} \leq c)\right]. \tag{4}$$

We have the following two lemmas (the proofs are straightforward):

Lemma 1 *Under the assumptions* (H_0) *and* (H_1)*, the process* $(X_t)_t$ *defined by equation (4) is locally second-order stationary (within each regime) and thus globally stationary. Its autocovariance function is :*

$$\gamma_X(h) = \frac{\Gamma(1-2d)\Gamma(h+d)}{\Gamma(d)\Gamma(1-d)\Gamma(h+1-d)} I_t(X_{t-1} \leq c) + \left[1 - I_t(X_{t-1} \leq c)\right], \quad (5)$$

where $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$, $a \geq 0$.

Lemma 2 *Under the assumptions* (H_0) *and* (H_1)*, the process* $(X_t)_t$ *defined by equation (4) has a spectral density function given by*

$$f_X(\omega) = \frac{1}{2\pi} \left|1 - e^{-i\omega}\right|^{-2d} I_t(X_{t-1} \leq c) + \frac{1}{2\pi} \left[1 - I_t(X_{t-1} \leq c)\right], \quad (6)$$

where $\omega \in [0, \pi]$.

As it is seen from the above expressions, the autocovariance function and the spectrum density depend upon the regime-shift variable. The "mixture" of a white noise process and of a fractional white noise process produces a memory structure that is function of the distribution function of the variable X_{t-1} across the two regimes at different dates. If regime 2 is more frequently visited by the observations than regime 1, then this will imply some difficulties to find a long-memory dynamics. In that case, the autocovariance and spectrum will exhibit a shape resembling to that of a short-memory process. In the opposite case, the autocovariance function and the spectrum exhibit the usual properties of long-memory processes : a slow decay and high values at frequencies near zero. The key parameter here is the threshold c, that determines the distribution function of the observations across the two regimes. Figure 1 shows different cases illustrating the memory properties of (2) for different values of the parameter d and different partitions of observations between the two regimes.

Asymptotically, the long-memory behavior dominates: the spectrum becomes infinite at the zero frequency and the autocovariance is not summable. We therefore have the following lemma:

Lemma 3 *Under the assumptions* (H_0) *and* (H_1)*, the asymptotic expressions of the autocovariance function and the spectral density function are respectively:*

$$\gamma_X(h) = \frac{\Gamma(1-2d)}{\Gamma(d)\Gamma(1-d)} h^{2d-1} + O(1), \quad \text{as } h \to +\infty \quad (7)$$

and

$$f_X(\omega) = C\omega^{-2d} + O(1), \quad as \quad \omega \to 0, \tag{8}$$

where C *is a positive constant.*

Although our concept of long-memory refers only to zero frequency, the above approach could be extended to generalized long-memory processes.

A simulation study is carried out in order to explore the memory properties of series that are generated according to equation (2). We examine the sample autocorrelation function and the periodogram. The simulation is set as follows.

1. In a first step, two series $(X_t^{(1)})_t$ and $(X_t^{(2)})_t$ are generated using the following equations:

$$X_t^{(1)} = (1-B)^{-d}\varepsilon_t, \quad X_t^{(2)} = \varepsilon_t, \quad \varepsilon_t \sim N(0,1), \quad t = 1, ..., T. \tag{9}$$

For purpose of illustration, we report below the results for the two sample sizes $T = 2000$ and $T = 10000$. We choose different values of d : $0.05, 0.1, 0.2, 0.3, 0.4, 0.49$.

2. In a second step, we define a method to partition the observations across the two regimes. An ideal way to proceed would be to derive the analytical expressions of the quantiles of the conditional distribution of a process such as (2) and use it to separate the two regimes. However, such a derivation is not immediate. Thus, we use the following alternative procedure. Define $X_{\max}^{(1)}$, $X_{\max}^{(2)}$ as the highest values of the series $X_t^{(i)}, i = 1, 2$ and $X_{\min}^{(1)}$, $X_{\min}^{(2)}$ as their lowest values. Further, consider the interval

$$[a_1, a_2] = \left[\max\left\{X_{\min}^{(1)}, X_{\min}^{(2)}\right\}, \ \min\left\{X_{\max}^{(1)}, X_{\max}^{(2)}\right\}\right].$$

The values of c are defined as percentages of the spread of this interval, that is $c_k = (a_1 - a_2)\, k + a_2$, $k = 10\%,\ 20\%,\ ...95\%$. When $k = 50\%$, we approximately have half of the observations within each regime. When k is increased, the number of observations in regime 1 decreases. For instance, if $k = 90\%$, $c_k = 0.9a_1 + 0.1a_2$; therefore, c_k will be relatively low and a few observations will belong to regime 1, following (2).

3. For a given c_k, we create a vector X_t in which the observations are stored as follows. The first observation X_1 is chosen randomly from $X_1^{(1)}$ and

$X_1^{(2)}$. The second observation is defined as $X_2 = X_2^{(1)}$, if $X_1 \leq c_k$, and as $X_2 = X_2^{(2)}$, if $X_1 > c_k$. The next observations of X_t are defined similarly.

4. We estimate the values of the long-memory parameter for the processes $(X_t)_t$, $(X_t^{(1)})_t$, $(X_t^{(2)})_t$ using the Geweke and Porter-Hudak (1983) estimator, the Robinson (1994) Gaussian semiparametric estimator and the Lobato and Robinson (1996) average periodogram estimator.

5. The above steps are repeated S times ($S = 100$, 500, 1000) and we draw the empirical distribution of the different estimators using an Epachenikov kernel. Finally, we report the value of the long-memory parameter corresponding to the mode of the distribution.

In tables 1 and 2 (to avoid too many tables, we report a selection of our results; other results are available upon request), we give the results of the estimations obtained for d. In these tables, we distinguish the results with respect to the percentage of the points inside the first regime. In table 1, the results are obtained using a sample size of $T = 10000$ and $S = 500$ replications, and in table 2, we use $T = 2000$ and $S = 100$.

The values of the estimated parameter d given in the tables are very suggestive. The mixture of the two regimes implies that, very often, the parameter d estimated using the whole process $(X_t)_t$ is lower than the parameter used to simulate the data in the first regime. For a given d, the fractional white noise regime is detected, only if they are "enough" data in the first regime (here for instance this regime is more easily detected when $k = 0.75$ than when $k = 0.25$). So, the important point is the percentage of observations in one regime in comparison to the other. Also, it is seen that the smaller the value of the long-memory parameter, the more predominant the white noise regime.

3 Locating the threshold parameter

3.1 General methodology

In practice, the parameter c is unknown. This parameter plays a crucial role in determining the way the observations are distributed across the two regimes. Different approaches are used in this paper. One is nonparametric and based on the properties of the spectral density function. The others are parametric and based on the minimization of the sum of squared residuals or on the $t-$statistics of the estimation of d. We use Monte Carlo procedures to examine the location problem of the parameter c. In the sequel, we shall assume that the threshold variable is known.

Consider a time series $(X_t)_t$ and assume that X_{t-1} is the threshold variable. Denote $(\widetilde{X_t})_t$ the arranged time series according to the decreasing values of X_{t-1}. Finding the threshold parameter c on the process $(\widetilde{X_t})_t$ rather than on the process $(X_t)_t$ allows us to consider the $SETAR$ model as a change-point problem. The beginning observations of the process $(\widetilde{X_t})_t$ are in the white noise regime, while the last observations are in the fractional white noise regime because the values of X_{t-1} are ranked in a decreasing order. The unknown parameter c delimitates the two regimes, somewhere in the 'middle' - not necessarily the half - of the vector $(\widetilde{X_t})_t$. All the methods described in the next section rely upon procedures that amount to study a succession of "local" - short and long - memory structures. One begins with a window of observations with a given length, that controls the initial number of observations. The latter are used to compute different indicators of long-memory (these can be estimates for the parameter d, test ratios using the spectral estimates, the Hurst statistic, ...). A fixed number of observations is then added to the initial vector and the indicators are recomputed, and so on. Since the initial observations in $(\widetilde{X_t})_t$ are in the white noise regime and the last observations are in the fractional white noise regime, one expects to detect a modification in the degree of persistence of the time series where new observations are added to the series. Ideally, we would expect to find a breakpoint indicating the threshold value at which the persistence structure in the memory changes. In practice, we face two problems. The first one is classical to all $SETAR$ processes, in the sense that the transition from one regime to another is continuous. A standard approach to overcome this difficulty is to argue that the threshold that is detected is solely an indication of a region where a partition occurs. One accordingly needs to try several values and compare the estimated models using different criteria (for instance, some information criteria, tests based on forecasts,...). The second problem is more specific to the type of $SETAR$ process considered here and will be discussed in more detail. Indeed, in regard to the results of the simulation in section 2, the finding of a long-memory structure in a vector of observations that consists of observations that are a mixture of - short and long - memory processes, requires that the long-memory regime contains enough observations. The implication here is that the number of observations for which a threshold is detected, necessarily overestimates the number of observations corresponding to the true - but unknown - threshold parameter. Clearly, such an overestimate implies a bias in the value of c, which can be either positive or negative. The sign of the bias can be determined via Monte Carlo experiments (see the next sections).

3.2 An approach based on spectral estimates

When the observations are in regime 1, the spectral density tends to infinity for frequencies near 0, while this is not the case for the white noise regime. We can use this property as an indication of the degree of memory in the data. The procedures works as follows.

1.- Choose an initial number of points t_1 in $(\widetilde{X}_t)_t$ and a finite number of frequencies near 0, denoted $\omega_1, \cdots, \omega_n$, where n is a small integer (in our case $n = 10$). Evaluate the sample version of the spectral density function at these frequencies and compute their mean

$$f_{t_1}(\omega) = \sum_{i=1}^{n} f_{t_1}(\omega_i). \tag{10}$$

2.- Repeat step 1, by augmenting the number of initial observations, while fixing the number of frequencies to n. One obtains a sequence $f_{t_1}(\omega)$, $f_{t_2}(\omega)$, ...

3.- Draw a scatterplot of this sequence versus t_1, t_2, ... The presence of a $SETAR$ model likewise equation (2) must be indicated by an inverse L curve, with a graph which is flat while the observations are in the white noise regime and a seemingly vertical line when one enters in the fractional white noise regime.

4. To locate the set of observations for which a changing dynamics - in terms of persistence - exists, one needs to implement a test. Here, we apply the Lobato and Robinson (1998)'s procedure using the n spectral densities estimated at step 2. For a given t_i, the statistic used is

$$LR_{t_i} = -\sqrt{n}\frac{\sum_{j=1}^{n} \upsilon_j I(\omega_j)}{\sum_{j=1}^{n} I(\omega_j)}, \quad \text{where} \quad \upsilon_j = \log j - n^{-1}\sum_{j=1}^{n} j; \tag{11}$$

$I(\omega_j)$ is an estimation of the periodogram associated to the process $(\widetilde{X}_t)_t$. This statistic is distributed as $N(0,1)$ variate. One thus has a sequence of statistics LR_{t_1}, LR_{t_2}, ...

5. Select the observation of the vector $\widetilde{X}_t$ that corresponds to the first statistic - in the sequence LR_{t_1}, LR_{t_2}, ... - with a significance level less than a given nominal size (generally 5% or 10%). Note that $\widetilde{c}$, the threshold value, is a biased estimate of the true value c, due to the number of "excess" observations needed to detect the long-memory regime.

To obtain an estimate of the excess number of observations, one need to draw simulations. This is done by repeating steps 1-5, S times ($S = 100, 500, 1000$). Here, we show the results for $S = 100$. Using one thousand replications instead of one hundred yields similar results, but it is computationally more time consuming in the second case.

Figure 2 shows an example of the scatterplot of the sequence of the mean spectrum computed as indicated in steps 1- 2. The vector containing a mixture of the two regimes is constructed in a similar way as indicated in section 2. We consider different values of the parameter d to simulate the process $(X_t)_t^{(1)}$ which contains the fractional white noise observations. The inverse L curve is very suggestive on the graph.

In table 3, the column labeled (7) indicates the values of the overestimate of the true number of observations in the white noise regime and the implication on the bias of c. The numbers in the table are the modes of the empirical distribution over 100 replications. The results suggest that, in general, the presence of a long-memory regime is detected when there are twice more observations than the number of effective observations in regime 2. And it is seen that this induces a negative bias on c. Consequently, in empirical applications, the location of a threshold point $\widehat{c}$ can serve as a benchmark below which the true value c has to be searched. In practice, one first locates a value $\widehat{c}$, then estimates different models with $c < \widehat{c}$ and finally retains the best model according to different criteria : $RMSE$, forecasting performance,... The choice of the different values for c approximately corresponds to regions where the arranged vector of observations contains half less observations than the number of observations corresponding to the located value $\widehat{c}$.

3.3 A grid search method

Let $[\underline{c}, \overline{c}]$ be the interval of plausible threshold values and consider $(\widetilde{X}_t)_t$ the time series of ordered observations defined above. Define the set of threshold candidates as

$$\left\{c_{(1)},\ c_{(2)}, \cdots, c_{(k)}, \cdots, c_{(n)}\right\} \subset \left[\underline{c}, \overline{c}\right],$$

with $c_{(k+1)} - c_{(k)} = \lambda$. The choice of λ depends upon the trade-off between computation time and expected precision. In practice, the candidates are chosen as the percentiles of the distribution function associated to the process $(\widetilde{X}_t)_t$. The grid search approach consists of the following steps.

1.- Partition the observations in $(\widetilde{X}_t)_t$ according to $c_{(1)}$, in order to obtain two series $(\widetilde{X}_t^{(1)})_t$ and $(\widetilde{X}_t^{(2)})_t$.

2.- Estimate $\widehat{d}^{(1)}$ and $\widehat{d}^{(2)}$ the long-memory parameters of the two series and compute the sum of squares residuals $SSR_{(1)}$ of the model: $\forall t$,

$$\begin{cases} (1-B)^{\widehat{d}^{(1)}} \widetilde{X}_t^{(1)} = \widehat{\varepsilon}_t^{(1)}, & \text{regime 1} \\ (1-B)^{\widehat{d}^{(2)}} \widetilde{X}_t^{(2)} = \widehat{\varepsilon}_t^{(2)}, & \text{regime 2.} \end{cases} \tag{12}$$

3.- Repeat steps 1-2 using $c_{(2)},\ c_{(3)},\ \cdots,\ c_{(n)}$. Let $SSR_{(2)}$, $SSR_{(3)}$, ..., $SSR_{(n)}$ the corresponding sum of the squares residuals.

4.- Finally, retain the lowest sum of squares residuals. Denote $\widetilde{c}$ the value of c corresponding to this sum.

Similarly to the approach based on spectral estimates, one may try to evaluate the performance of the grid method by drawing the distribution of the bias via Monte Carlo simulations. Figures (3a) to (3d) show the distribution of the bias $(\widehat{c}-c)$ for a sample length $T = 500$ and $S = 100$ replications of the method. The initial vector of observations $(X_t)_t$ is generated using different values of d. For purpose of illustration, we have selected here two cases corresponding to $d = 0.49$ and $d = 0.10$. We use two estimators : the GPH and the Robinson semiparametric estimators. As it is seen, the distributions are bimodal, thereby showing both positive and negative biases of the parameter c. The simulations showed that a positive bias was a consequence of an underestimate of the number of observations in regime 2, while a negative bias was induced by an overestimate. However, a positive bias indicates a spurious changing memory structure. Indeed, detecting a threshold value when the number of observations in regime 2 is less than the true number of observations, would mean that one finds a long-memory structure in the white noise regime! Consequently, the bias must necessarily be negative. The detection of a spurious changing memory structure can be explained as follows. In the white noise regime (regime 2), we occasionally found estimates of the long-memory parameter that were negative and statistically significant. As one knows, it is an open problem as whether the distinction between a white noise process and a long-memory process with anti-persistence property is possible. Sample estimates of their density functions exhibit a similar shape. In our case, when the parameter $\widehat{d}^{(2)}$ is negative, this increases the squares residuals $\sum \left(\widehat{\varepsilon}_t^{(1)}\right)^2$ and it is possible to find values of the sum of squares residuals that are minimized at some threshold point $\widehat{c} > c$.

Using the sum of squares residuals as a criterion for selecting the threshold values from the grid search method is thus inadequate and this is also true for criteria such as AIC, BIC or the variance.

3.4 Locating change-points in the t-ratios

We consider the time series $(\widetilde{X}_t)_t$ of arranged observations according to the decreasing values of X_{t-1}. The change-point problem using the t-ratio of the long-memory parameter is formulated as follows.

1. One considers a set of s_1 initial observations of $(\widetilde{X}_t)_t$ and estimates the long-memory parameter and the corresponding $t - ratio$, t_{s_1}.

2. The vector $(\widetilde{X}_t)_t$ is incremented in such a way to contain s_2, s_3, ..., s_n observations and new $t - ratios$ are computed: t_{s_2}, t_{s_3},... , t_{s_n}.

3. Consider the set of estimated t-ratios $\{t_{s_1}, t_{s_2}, t_{s_3}, ..., t_{s_n}\}$. One tests the presence of a structural break $\widetilde{t}$ in the view of finding a sequence of t-ratios such that, for $t_{s_k} \leq \widetilde{t}$, the estimated long-memory parameters are not statistically significant, while, for $t_{s_k} > \widetilde{t}$, they are significant.

A simple way to implement step 3 is to use a standard Chow test. The series of t-ratios is regressed on a linear time trend, using incremented dummy variables: for $k = 1, 2, \ldots, n$

$$t_{s_{k_t}} = (a + bD_t) + (c + dD_t)t + u_t, \tag{13}$$

with $u_t \sim iid$ and

$$D_t = \begin{cases} 1, & if\ t \leq \widetilde{t} \\ 0, & otherwise \end{cases}$$

and we test the null hypothesis $H_0 : b = d = 0$ against $H_1 : b \neq 0$ or $d \neq 0$.

The constant term is omitted if we want to test changes only in the slope. The test is implemented by considering different values of $\widetilde{t}$ and retaining finally the value yielding the lowest p-value.

To evaluate the adequacy of such a method for the delimitation of the two regimes, we conduct Monte Carlo simulations on series of length $T = 500$ with $S = 100$ replications. The original series $(X_t)_t$ is generated according to equation (2) for different values of the parameter d. We show, in table 3, a selection of results (other results are available upon request to authors). More precisely, we distinguish between three cases:

(i) $a \neq 0$ and $b \neq 0$, i.e. we consider (13) and $H_0 : b = d = 0$ against $H_1 : b \neq 0$ or $d \neq 0$;

(ii) $a \neq 0$ and $b = 0$, i.e.

$$t_{s_{k_t}} = a + (c + dD_t)t + u_t \tag{14}$$

and $H_0 : d = 0$ against $H_1 : d \neq 0$;

(iii) $a = 0$ and $b = 0$, i.e.

$$t_{s_{k_t}} = (c + dD_t)t + u_t \tag{15}$$

and $H_0 : d = 0$ against $H_1 : d \neq 0$.

Table 3 reports the mode of the empirical distribution of this difference over 100 replications of the tests, along with the percentage of excess observations needed in regime 2 to detect a changing dynamic in the memory behavior of the time series (see the columns labeled (1)-(2)-(3)). The result are similar to those of the spectral approach, in the sense that an overestimate of the number of observations in regime 2 induces a negative bias on the estimate of the parameter c. Consequently, similar recommendations concerning the empirical applications apply here. A difference is, however, that the overestimate of the true number of observations is less important in comparison to the approach based on spectral estimation. Figures (4a) to (4f) show examples of the distribution of the difference $(\widehat{c} - c)$ for $d = 0.40$.

Instead of using the Chow test, one can also compute the sum of squares residuals corresponding to equations (13), (14) or (15), and select the t-ratio (and thus the threshold value) yielding the lowest sum. We consider the same three cases as before; the results are shown in table 3, in the columns labeled (4), (5) and (6). Clearly, the true value is again underestimate, but it is seen that in some cases (in particular in the case (*iii*): $a = b = 0$) the bias is very low. Figure (5a) to (5f) shows an example of the sample distribution of the bias for $d = 0.49$.

4 Empirical applications

We consider different asset prices and five stock indices. The data are daily and run from 1997 to 2003. The names of the companies and indices are listed in table 4.

It is a well documented stylized fact that there is little serial correlation in the levels of returns, when the latter are computed by considering changes between two consecutive prices. Also, a recent result by Granger (2002), that contradicts many conclusions obtained henceforth in the literature, suggests that absolute returns and squared returns cannot have long-memory properties of the $ARFIMA$ type. Consequently, it seems difficult to apply our $SETAR$ model to measures of returns and volatility that are usually considered in the literature. For both the standard logarithmic returns and volatility measures (absolute or squared returns), the presence of regime 1 is excluded. Other measures of returns are, however, capable of producing the dynamics inherent to equation (2). Remembering the arguments concerning the implications of the presence of transaction costs on financial markets (see the introduction), the horizons of intervention on the markets are likely not to be daily. Rather, we can imagine, for instance, that the positions are periodically revised at the end of a trading week. We accordingly consider the following measure of the returns:

$$R_{t,t-j} = \frac{\log(p_t) - \log(p_{t-j})}{j}, \quad 0 < j < t, \tag{16}$$

where p_t is the stock price or index at time t and $j = 1, 2, 3, ...$ On daily prices quotes, $R_{t,t-j}$ is the logarithmic return between days t and $t - j$. The longer the period separating two prices, the more likely the no-arbitrage condition not to hold (because prices changes at time t implies delayed reactions from the agents) and the higher the likelihood to find a long-memory structure in the data.

In table 5, we report the estimated values of d for different j for all the series presented in table 4. Clearly, for $j = 1$, the long-memory hypothesis is rejected, while it is generally accepted for $j > 1$. The degree of persistence increases with the value of j. For $j > 4$, the estimated values of d were higher than 0.5, thereby signaling a nonstationary dynamics in the returns. Instead of considering all the values of j to apply our SETAR model, we stick to one case ($j = 3$) chosen for purpose of illustration.

Table 6 contains the estimates of the parameters d_1 and d_2 corresponding to regimes 1 and 2, using the procedures exposed in the preceding section. A long memory behavior in one regime is found for a majority of series. We observe that the nonparametric approach yields estimates of the long-memory parameter $\widehat{d_1}$ that are close to those of the parametric methods for half of the series. When the results are different, the spectral approach often yields values of $\widehat{d_1}$ that are higher. This can indicate an overestimate of the true parameter (although we tried more than 50 models by considering different values of the threshold below the "breakpoint" indicated on the graph of the mean spectrum). Differences in the estimations between the parametric and nonparametric approaches come from the number of observations detected within each regime.

For purpose of comparison, we use in-sample observations to compare the forecastability performance of the different models. In this view, we use a battery of tests based on loss functions: the asymptotic test, the sign tests, then Wilcoxon's test, the Naive benchmark test, The Morgan-Granger-Newbold test and the Meese Rogoff test (all these tests are summarized in Diebold and Mariano (1995)). We compare the model based on the spectral approach and the model based on the parametric method (the model with the lowest RMSE). The p-values in table 7 clearly indicate that the null hypothesis of equal accuracy of both models is strongly rejected. The last column gives the number of times the model based on the parametric approach outperforms the model estimated using the spectral approach. For some series the parametric model is more performing, while for other series the spectral based approach dominates.

5 Conclusion

This paper has proposed a new model to investigate the long-memory dynamics of time series that contains a mixture of long-memory and white noise structures. An important question concerns the global behavior of a model such as (2). We want to use it to detect existence of long memory behavior and its shift between the two regimes. We observe, by simulations, that the shift from global short memory behavior to global long memory behavior for such a model depends strongly upon the threshold c or from the percentage of observed points inside one regime or another. It will be interesting to understand deeper this phenomena which is probably the key for a good identification of these models. This will do in another paper. To pick up the part of the long-memory in the data, we suggest both nonparametric and parametric methods in order to locate the threshold parameter. We show that the methods fit the data well, when the returns are computed by considering the presence of a delay between price changes.

Other further research topics are the following. Firstly, one can conjecture that the model suggested here, can be extended to the case where regime 1 is described by an $ARFIMA$ model and regime 2 by a stationary $ARMA$ model or by a mixing process. Such a question is interesting since an ARFIMA process is not mixing (see Guégan and Ladoucette (2001)). Secondly, the model can be applied to time series other than returns. Variables based on technical trading rules can be suggested (for instance, a variable constructed from a short-run moving average and a long-run moving average) in order to capture asymmetric dynamics in the memory structure. Thirdly, the $SETAR$ model with long-memory behavior can be studied under the assumption of heteroskedastic errors in order to incorporate the influence of volatility of the long-memory structure. Finally, the identification and forecasting problems are not examined here, but it would be interesting to compare this model to others producing persistent dynamics, such as switching processes and the class of all long-memory models.

Acknowledgments

This paper was prepared for the International Conference on Forecasting Financial Markets, 4-6 June 2003, in Paris, France. The authors thank the Conseil Régional of PACA for its financial support. We also thank Aldric Pouron for providing us with the data.

References

[1] Bos, C., Franses, P. H., and Ooms, M. (1999), "Re-analyzing inflation rates : evidence of long-memory and level shifts", *Empirical Economics,* 24, 427-449.

[2] Clarida, R. H., and Taylor, M. P. (2001), "Nonlinear permanent-temporary decomposition, with applications in macroeconomics and finance", Working Paper, University of Warwick.

[3] Diebold, F. X., and Inoue, A. (2001), "Long-memory and regime switching", *Journal of Econometrics,* 105, 131-159.

[4] Diebold, F. X., and Mariano, R. (1995), "Comparing predictive accuracy", *Journal of Business and Economic Statistics*, 253-263.

[5] Dufrénot, G., and Mignon, V. (2002), *Recent developments in nonlinear cointegration with applications to macroeconomics and finance*, Kluwer academic Press, Boston.

[6] Dufrénot, G., Mignon, V., and Péguin-Feissolle, A. (2003), "Business cycles asymmetry and monetary policy: A further investigation using MRSTAR models", *Economic Modelling,* in press.

[7] Franses, P. H., and Paap, R. (2002), "Censored latent effects autoregression, with an application to US unemployment", *Journal of Applied Econometrics,* 17, 347-366.

[8] Franses, P. H., van Der Leij, M., and Paap, R. (2002), "Modelling and forecasting level shifts in absolute returns", *Journal of Applied Econometrics,* 17, 601-616.

[9] Geweke, J., and Porter-Hudak, S. (1983), "The estimation and application of long-memory time series", *Journal of Time Series Analysis,* 4, 221-238.

[10] Granger, C. W. J. (2002), "Long-memory, volatility, risk and distribution", Discussion paper, University of California, San Diego.

[11] Granger, C. W. J., and Hyung, N. (1999), "Occasional breaks and long-memory", Discussion paper 99-14, University of California, San Diego.

[12] Guégan D. (2000), "A new model: the k-Factor GIGARCH Process", *Journal of Signal Processing*, 4, 265 - 271.

[13] Guégan D. (2003), "A Prospective study of the k-Factor Gegenbauer Processes with heteroscedastic errors and an application to inflation rates", *Finance India*, XVII, 165-197.

[14] Guégan D., and Ladoucette S. (2001), "Non mixing properties of long memory processes", *C.R.A.S.* 333, Ser. I, 373-376.

[15] Hidalgo, J., and Robinson, P. M. (1996), "Testing for structural change in a long-memory environment", *Journal of Econometrics,* 70, 159-174.

[16] Lobato, I. N., and Robinson, P. M. (1998), "A nonparametric test for $I(0)$", *Review of Economic Studies,* 65, 475-495.

[17] Lobato, I. N., and Savin, N. E. (1997), "Real and spurious long-memory properties of stock-market data", *Journal of Business and Economic Statistics,* 16, 261-283.

[18] Martens, M., Kofman, P., and Vorst, T. C. (1998), "A threshold error-correction model for intraday futures and index returns", *Journal of Applied Econometrics,* 13, 245-263.

[19] Robinson, P. M. (1994), "Semiparametric analysis of long-memory time series", *Annals of Statistics,* 22, 515-539.

[20] Sarno, L., and Taylor, M. (2001), "Real exchange rate dynamics in transition economies: a nonlinear analysis", *Studies in Nonlinear Dynamics and Econometrics,* 5, 153-177.

[21] van Dijk, D., Franses, P. H., and Paap, R. (2002), "A nonlinear long-memory model with an application to US unemployment", *Journal of Econometrics,* 110, 135-165.

Table 1: Estimation of the long memory parameter d using three methods, with $T = 10000, S = 500$, with respect to the percentage of points in regime 1, for the process (4) (T-ratios in parentheses)

	value of d		% obs. in reg. 1	% obs. in reg. 1	% obs. in reg. 1
			75%	50%	25%
GPH	0.05	$\widehat{d}$	0.0628 (0.91)	0.0431 (0.65)	-0.0088 (-0.13)
		$\widehat{d}_1$	0.0480 (0.67)	0.0244 (0.29)	0.0586 (0.32)
		$\widehat{d}_2$	0.0467 (0.26)	-0.0065 (0.07)	0.0118 (0.16)
	0.30	$\widehat{d}$	0.3062 (4.57)	0.2551 (3.83)	0.0149 (0.22)
		$\widehat{d}_1$	0.2818 (4.04)	0.3125 (3.88)	0.1915 (1.10)
		$\widehat{d}_2$	0.0182 (0.11)	-0.0055 (-0.00)	0.0078 (0.10)
Semiparametric	0.05	$\widehat{d}$	0.0452 (3.60)	0.0152 (1.21)	0.0048 (0.38)
		$\widehat{d}_1$	0.0452 (3.53)	0.0354 (2.18)	0.0192 (0.54)
		$\widehat{d}_2$	0.0040 (0.07)	-0.0046 (-0.25)	0.0040 (0.35)
	0.30	$\widehat{d}$	0.2849 (22.68)	0.1956 (15.57)	0.0054 (0.43)
		$\widehat{d}_1$	0.2853 (22.30)	0.2746 (16.61)	0.1755 (3.16)
		$\widehat{d}_2$	-0.0094 (-0.27)	0.0043 (0.29)	0.0040 (0.33)
average periodogram	0.05	$\widehat{d}$	0.0515 (0.03)	0.0366 (0.02)	0.0227 (0.00)
		$\widehat{d}_1$	0.0473 (0.01)	0.0087 (0.00)	-0.0059 (-0.08)
		$\widehat{d}_2$	0.0448 (0.06)	0.0178 (0.03)	0.0082 (0.01)
	0.30	$\widehat{d}$	0.2791 (1.03)	0.2353 (0.23)	0.0094 (0.00)
		$\widehat{d}_1$	0.2779 (1.01)	0.2683 (0.20)	0.1862 (0.08)
		$\widehat{d}_2$	0.0559 (0.06)	0.0247 (0.03)	-0.0056 (-0.02)

Table 2: Estimation of the long memory parameter d using three methods, with $T = 2000, S = 100$, with respect to the percentage of points in regime 1, for the process (4) (T-ratios in parentheses)

	value of d		% obs. in reg. 1	% obs. in reg. 1	% obs. in reg. 1
			75%	50%	25%
GPH		$\widehat{d}$	0.0405 (1.26)	0.0210 (0.65)	0.0027 (0.08)
	0.05	$\widehat{d}_1$	0.0491 (1.48)	0.0155 (0.36)	-0.0002 (-0.00)
		$\widehat{d}_2$	0.0035 (0.02)	0.0012 (0.02)	-0.0048 (-0.14)
		$\widehat{d}$	0.1825 (5.75)	0.1152 (3.65)	0.0306 (0.96)
	0.20	$\widehat{d}_1$	0.1851 (5.58)	0.1427 (3.35)	0.1226 (1.01)
		$\widehat{d}_2$	-0.0098 (-0.09)	-0.0010 (-0.02)	0.0079 (0.24)
		$\widehat{d}$	0.3538 (11.05)	0.3032 (9.50)	0.0758 (2.37)
	0.40	$\widehat{d}_1$	0.3859 (11.50)	0.3452 (8.40)	0.3308 (3.01)
		$\widehat{d}_2$	0.0118 (0.12)	0.0121 (0.27)	0.0096 (0.29)
		$\widehat{d}$	0.4547 (14.19)	0.4097 (12.92)	0.1238 (3.97)
	0.49	$\widehat{d}_1$	0.4557 (13.71)	0.4595 (11.34)	0.4391 (5.15)
		$\widehat{d}_2$	0.0150 (0.16)	-0.0058 (-0.13)	0.0050 (0.15)
Semiparametric		$\widehat{d}$	0.0449 (1.88)	0.0245 (1.02)	-0.0010 (-0.04)
	0.05	$\widehat{d}_1$	0.0345 (1.41)	0.0207 (0.80)	0.0072 (0.13)
		$\widehat{d}_2$	-0.0138 (-0.02)	-0.0015 (-0.08)	-0.0008 (-0.05)
		$\widehat{d}$	0.1853 (7.74)	0.1150 (4.80)	0.0215 (0.89)
	0.20	$\widehat{d}_1$	0.1849 (7.57)	0.1492 (4.86)	0.0420 (0.53)
		$\widehat{d}_2$	-0.0247 (-0.04)	-0.0043 (-0.02)	0.0003 (0.00)
		$\widehat{d}$	0.3443 (14.39)	0.3055 (12.77)	0.0787 (3.29)
	0.40	$\widehat{d}_1$	0.3795 (14.96)	0.3461 (11.29)	0.3121 (4.39)
		$\widehat{d}_2$	-0.0044 (-0.19)	0.0092 (0.29)	0.0108 (0.43)
		$\widehat{d}$	0.4451 (18.61)	0.4163 (17.40)	0.1147 (4.79)
	0.49	$\widehat{d}_1$	0.4642 (18.84)	0.4496 (15.04)	0.4225 (6.39)
		$\widehat{d}_2$	-0.0088 (-0.08)	0.0077 (0.20)	0.0111 (0.44)
average periodogram		$\widehat{d}$	0.0426 (0.04)	0.0208 (0.01)	0.0105 (0.01)
	0.05	$\widehat{d}_1$	0.0446 (0.03)	0.0223 (0.01)	0.0527 (0.05)
		$\widehat{d}_2$	-0.0493 (-0.02)	0.0049 (0.00)	0.0098 (0.00)
		$\widehat{d}$	0.1653 (0.22)	0.1224 (0.15)	0.0292 (0.02)
	0.20	$\widehat{d}_1$	0.1825 (0.27)	0.1309 (0.15)	0.0718 (0.06)
		$\widehat{d}_2$	0.0104 (0.01)	0.0125 (0.01)	0.0124 (0.00)
		$\widehat{d}$	0.3257 (1.05)	0.2815 (0.75)	0.0751 (0.08)
	0.40	$\widehat{d}_1$	0.3399 (1.09)	0.3209 (1.02)	0.2815 (0.19)
		$\widehat{d}_2$	-0.0345 (-0.00)	-0.0097 (-0.01)	0.0002 (0.00)
		$\widehat{d}$	0.3872 (1.59)	0.3678 (1.20)	0.1433 (0.15)
	0.49	$\widehat{d}_1$	0.4036 (1.81)	0.3763 (1.56)	0.3327 (1.09)
		$\widehat{d}_2$	-0.0125 (-0.01)	0.0217 (0.02)	0.0001 (0.00)

Table 3: Percentage of excess observations in regime 2 for the process (4) and estimated bias $\widehat{c}-c$ (for the threshold) by GPH method, when $T = 500$ and $S = 100$

d	mode	(1)	(2)	(3)	(4)	(5)	(6)	(7)
0.49	$\widehat{c}-c$	-0.26	-0.18	-0.22	-0.10	-0.24	-0.16	-1.05
	% excess	17	15	23	12	17	16	47
0.40	$\widehat{c}-c$	-0.24	-0.10	-0.14	-0.11	-0.18	-0.18	-1.52
	% excess	19	23	16	13	19	20	52
0.20	$\widehat{c}-c$	0.20	-0.14	-0.10	0.09	-0.03	0.02	-0.51
	% excess	16	18	11	1	24	9	52
0.10	$\widehat{c}-c$	-0.13	-0.23	0.11	0.19	-0.22	0.10	-0.2
	% excess	12	16	5	-1	15	7	53
0.05	$\widehat{c}-c$	-0.03	-0.11	0.20	-0.14	-0.01	0.18	-0.15
	% excess	10	15	13	14	12	5	48

Note: The different methods are:

* Locating change-points in the t-ratios by using Chow test on
  (1) $t_{s_{k_t}} = (a + bD_t) + (c + dD_t)t + u_t$
  (2) $t_{s_{k_t}} = a + (c + dD_t)t + u_t$
  (3) $t_{s_{k_t}} = (c + dD_t)t + u_t$
* Locating change-points in the t-ratios by minimizing the sum of squared residuals:
  (4) $t_{s_{k_t}} = (a + bD_t) + (c + dD_t)t + u_t$
  (5) $t_{s_{k_t}} = a + (c + dD_t)t + u_t$
  (6) $t_{s_{k_t}} = (c + dD_t)t + u_t$
* (7) Spectral estimation.

Table 4: Names of companies and indices with the period of estimation

N	Companies or indices	period
1	Bnp	02/01/98-23/04/03
2	Carrefour	02/01/98-23/04/03
3	Loreal	02/01/98-23/04/03
4	St Gobain	02/01/98-23/04/03
5	Total FinaElf	02/01/98-23/04/03
6	Vivendi	02/01/98-23/04/03
7	France Télécom	20/10/97-23/04/03
8	Lvmh	02/01/98-23/04/03
9	Ftse	02/01/98-23/04/03
10	Cac40	01/05/98-23/04/03
11	Sbf	02/01/98-23/04/03
12	Nasdaq 100	02/01/98-23/04/03
13	Dow Jones	02/01/98-23/04/03

Table 5: Estimation of d for different j using GPH estimator (the t-ratios are in parentheses), for prices and stocks

FRANCETEL			BNP			LVMH		
$j=1$	-0.0068	(-0.16)	$j=1$	-0.0548	(-1.54)	$j=1$	0.0210	(0.55)
$j=2$	0.0612	(1.61)	$j=2$	0.0070	(0.19)	$j=2$	0.1049	(2.52)
$j=3$	0.2073	(5.37)	$j=3$	0.1495	(3.82)	$j=3$	0.2349	(5.82)
$j=4$	0.4805	(10.43)	$j=4$	0.4320	(9.47)	$j=4$	0.5028	(10.27)
CARREFOUR			LOREAL			ST GOBAIN		
$j=1$	-0.0510	(-1.36)	$j=1$	-0.1341	(-3.58)	$j=1$	-0.0620	(-0.48)
$j=2$	0.0330	(0.80)	$j=2$	-0.0706	(-1.73)	$j=2$	0.0660	(1.77)
$j=3$	0.1294	(3.35)	$j=3$	0.0595	(1.49)	$j=3$	0.2054	(5.48)
$j=4$	0.3981	(8.46)	$j=4$	0.3587	(7.08)	$j=4$	0.5025	(10.38)
TOTAL			VIVENDI			FTSE		
$j=1$	-0.1308	(-3.51)	$j=1$	-0.0197	(-0.53)	$j=1$	-0.0874	(-2.21)
$j=2$	-0.0490	(-1.24)	$j=2$	0.0617	(1.44)	$j=2$	-0.0341	(-0.88)
$j=3$	0.0567	(1.54)	$j=3$	0.1903	(4.75)	$j=3$	0.1026	(2.69)
$j=4$	0.3260	(7.06)	$j=4$	0.4628	(9.30)	$j=4$	0.3587	(7.84)
CAC40			SBF			NASDAQ		
$j=1$	-0.0670	(-1.68)	$j=1$	-0.0419	(-0.94)	$j=1$	-0.0524	(-1.39)
$j=2$	0.0009	(0.02)	$j=2$	0.0145	(0.35)	$j=2$	0.0229	(0.60)
$j=3$	0.1338	(3.25)	$j=3$	0.1494	(3.46)	$j=3$	0.1649	(4.08)
$j=4$	0.4265	(8.40)	$j=4$	0.4139	(8.30)	$j=4$	0.4348	(9.45)
DOWJONES								
$j=1$	-0.0013	(-0.40)						
$j=2$	0.0814	(2.13)						
$j=3$	0.2057	(5.47)						
$j=4$	0.4779	(10.03)						

Table 6: Application to stock indices and asset prices for j=3 with different methods: (1) Locating change-points in the t-ratios using Chow test, (2) Locating change-points in the t-ratios minimizing the sum of squared residuals, (3) Spectral estimation

FRANCETEL

	T=1390	min=-0.1220		max=0.1110	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	0.0210	0.0211	0.0210	0.0108	0.0182
n_1	1190	1193	1190	990	1135
$\widehat{d}_1$	0.16	0.16	0.16	0.16	0.09
t-stat	(4.44)	(5.44)	(4.44)	(5.04)	(2.99)
n_2	200	197	200	400	255
$\widehat{d}_2$	-0.04	0.02	-0.04	-0.01	0.07
t-stat	(-0.53)	(0.32)	(-0.53)	(-0.21)	(1.24)

BNP

	T=1339	min=-0.0829		max=0.0625	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	0.0066	0.0123	0.0066	0.0097	-0.0157
n_1	917	1110	917	1033	162
$\widehat{d}_1$	0.10	0.11	0.10	0.12	0.12
t-stat	(2.49)	(3.63)	(2.49)	(3.84)	(1.81)
n_2	422	229	422	306	1177
$\widehat{d}_2$	0.09	0.00	0.09	0.08	0.01
t-stat	(1.62)	(0.00)	(1.62)	(1.57)	(0.33)

LVMH

	T=1340	min=-0.0811		max=0.0720	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	0.0220	0.0244	0.0220	0.0207	-0.0200
n_1	1230	1255	1230	1213	108
$\widehat{d}_1$	0.13	0.14	0.13	0.13	0.32
t-stat	(3.27)	(4.85)	(3.27)	(4.45)	(4.09)
n_2	110	85	110	127	1232
$\widehat{d}_2$	0.07	0.06	0.07	0.07	0.00
t-stat	(0.61)	(0.69)	(0.61)	(0.96)	(0.00)

CARREFOUR

	T=1339	min=-0.0710		max=0.0734	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	0.0112	0.0152	0.0112	0.0152	-0.0242
n_1	1114	1203	1114	1203	43
$\widehat{d}_1$	0.09	0.08	0.09	0.08	0.45
t-stat	(2.23)	(2.72)	(2.23)	(2.72)	(3.81)
n_2	225	136	225	136	1296
$\widehat{d}_2$	-0.10	-0.04	-0.10	-0.04	-0.02
t-stat	(-1.10)	(-0.56)	(-1.10)	(-0.56)	(0.69)

LOREAL

	T=1340	min=-0.0555		max=0.0656	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	0.0031	-0.0005	0.0031	-0.0005	0.0165
n_1	815	637	815	637	1200
$\widehat{d}_1$	0.04	0.01	0.04	0.01	0.06
t-stat	(0.91)	(0.26)	(0.91)	(0.26)	(2.04)
n_2	525	703	525	703	140
$\widehat{d}_2$	0.03	0.04	0.03	0.04	0.09
t-stat	(0.67)	(1.09)	(0.67)	(1.09)	(1.23)

ST GOBAIN

	T=1340	min=-0.0938		max=0.0531	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	-0.0076	0.0153	-0.0076	0.0113	-0.0027
n_1	352	1176	352	1093	527
$\widehat{d}_1$	0.17	0.19	0.17	0.17	0.08
t-stat	(2.29)	(6.41)	(2.29)	(5.57)	(1.95)
n_2	988	164	988	247	813
$\widehat{d}_2$	0.08	0.01	0.08	-0.03	0.04
t-stat	(1.91)	(0.15)	(1.91)	(-0.54)	(1.15)

TOTAL

	T=1340	min=-0.0588		max=0.0469	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	-0.0019	0.0054	-0.0019	0.0054	-0.0017
n_1	570	912	570	912	573
$\widehat{d}_1$	0.09	0.03	0.09	0.03	0.07
t-stat	(1.58)	(0.91)	(1.58)	(0.91)	(1.76)
n_2	770	428	770	428	767
$\widehat{d}_2$	0.01	-0.06	0.01	-0.06	-0.01
t-stat	(0.24)	(-1.35)	(0.24)	(-1.35)	(-0.28)

VIVENDI

	T=1338	min=-0.1787		max=0.1116	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	-0.0037	0.0062	-0.0037	0.0025	-0.0005
n_1	571	914	571	788	670
$\widehat{d}_1$	0.06	0.08	0.06	0.09	0.09
t-stat	(1.22)	(2.44)	(1.22)	(2.58)	(2.42)
n_2	767	424	767	550	668
$\widehat{d}_2$	0.04	0.03	0.04	-0.01	0.04
t-stat	(0.95)	(0.67)	(0.95)	(-0.24)	(1.06)

FTSE

	T=1338	min=-0.0359		max=0.0414	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	0.0007	-0.0004	0.0007	-0.0004	-0.0119
n_1	730	640	730	640	84
$\widehat{d}_1$	0.02	0.03	0.02	0.03	0.23
t-stat	(0.64)	(0.79)	(0.64)	(0.79)	(2.64)
n_2	608	698	608	698	1254
$\widehat{d}_2$	0.02	-0.02	0.02	-0.02	0.00
t-stat	(0.52)	(-0.54)	(0.52)	(-0.54)	(0.00)

CAC40

	T=1339	min=-0.0449		max=0.0547	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	-0.0009	0.0137	-0.0009	0.0137	0.0066
n_1	603	1243	603	1243	1026
$\widehat{d}_1$	0.10	0.12	0.10	0.12	0.06
t-stat	(1.85)	(4.14)	(1.85)	(4.14)	(1.91)
n_2	736	96	736	96	313
$\widehat{d}_2$	-0.03	-0.10	-0.03	-0.10	0.06
t-stat	(-0.59)	(-1.23)	(-0.59)	(-1.23)	(1.16)

SBF

	T=1340	min=-0.0401		max=0.0504	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	0.0055	0.0053	0.0055	0.00539	0.0002
n_1	980	971	980	971	676
$\widehat{d}_1$	0.09	0.11	0.09	0.11	0.10
t-stat	(2.28)	(3.44)	(2.28)	(3.44)	(2.69)
n_2	360	369	360	369	664
$\widehat{d}_2$	0.01	-0.06	0.01	-0.06	-0.05
t-stat	(0.16)	(-1.27)	(0.16)	(-1.27)	(-1.32)

NASDAQ

	T=1331	min=-0.0661		max=0.0562	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	0.0047	0.0209	0.0047	0.0209	0.0072
n_1	798	1232	798	1232	882
$\widehat{d}_1$	0.05	0.13	0.05	0.13	0.08
t-stat	(1.11)	(4.47)	(1.11)	(4.47)	(2.40)
n_2	533	99	533	99	449
$\widehat{d}_2$	0.00	0.00	0.00	0.00	0.06
t-stat	(0.05)	(0.00)	(0.05)	(0.0000)	(1.37)

DOWJONES

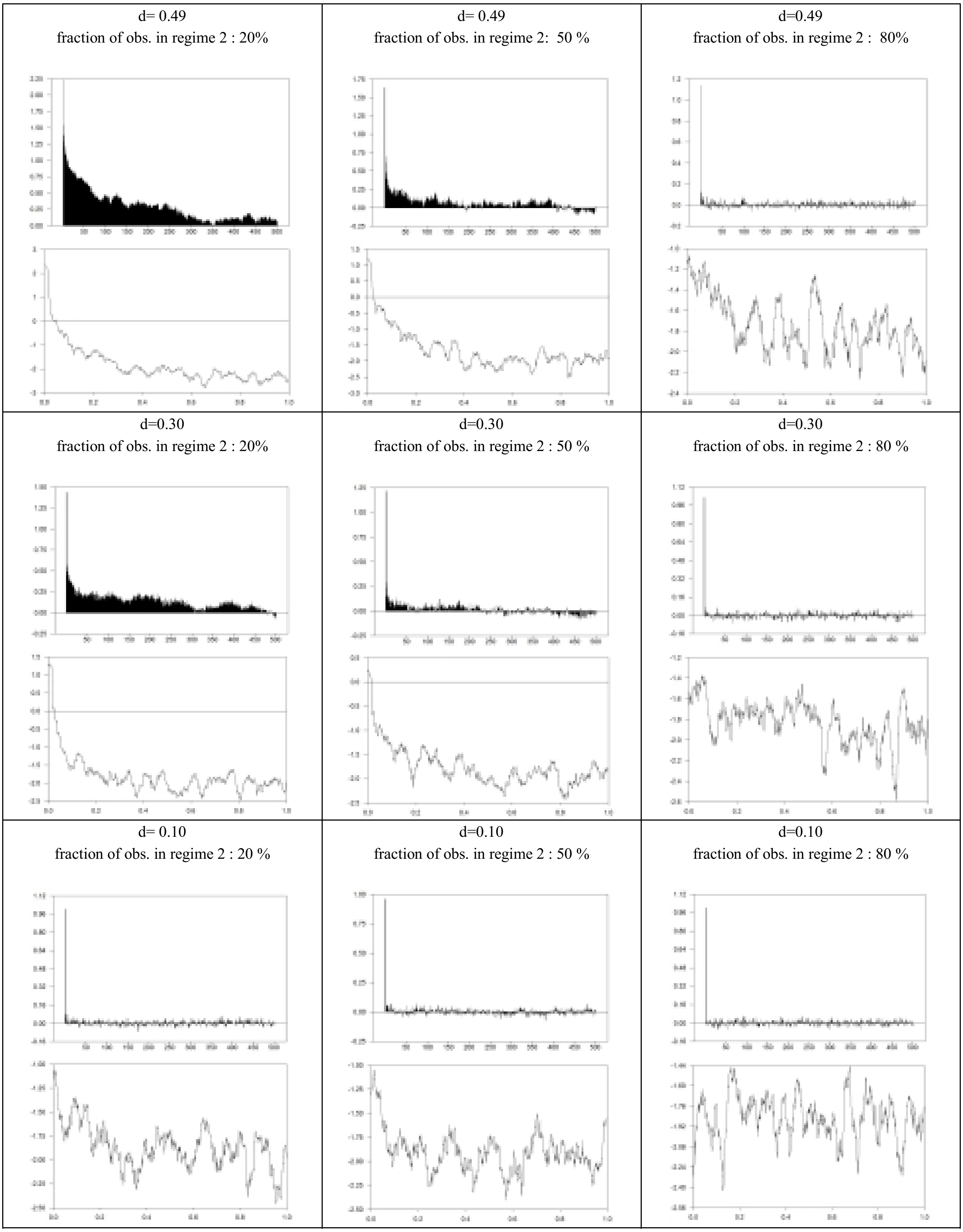
	T=1331	min=-0.0409		max=0.0304	
j=3	(1)	(1)	(2)	(2)	(3)
	GPH	SEMI-PAR	GPH	SEMI-PAR	GPH
$\widehat{c}$	0.0052	-0.0079	0.0052	-0.0079	-0.0072
n_1	1037	171	1037	171	191
$\widehat{d}_1$	0.12	0.18	0.12	0.18	0.15
t-stat	(3.00)	(2.81)	(3.00)	(2.81)	(2.41)
n_2	294	1160	294	1160	1140
$\widehat{d}_2$	0.02	0.08	0.02	0.08	0.00
t-stat	(0.33)	(2.68)	(0.33)	(2.68)	(0.00)

Table 7: Tests for comparing predictive accuracy - p-values

	AS	SI	WI	NB	MGN	MR	n
FRANCETEL	0.000	0.000	0.000	0.181	0.000	0.000	1.24
BNP	0.020	0.000	0.007	0.365	0.004	0.035	1.01
LVMH	0.241	0.286	0.000	0.328	0.064	0.162	1.31
CARREFOUR	0.686	0.547	0.000	0.435	0.507	0.596	1.30
LOREAL	0.000	0.000	0.000	0.265	0.000	0.000	0.72
STGOBAIN	0.075	0.000	0.324	0.356	0.005	0.030	1.10
TOTAL	0.000	0.000	0.000	0.090	0.000	0.000	0.46
VIVENDI	0.001	0.000	0.000	0.467	0.028	0.055	0.67
FTSE	0.000	0.000	0.000	0.027	0.000	0.000	0.23
CAC40	0.000	0.000	0.000	0.203	0.000	0.000	1.49
SBF	0.000	0.000	0.062	0.448	0.000	0.009	1.49
NASDAQ	0.000	0.227	0.000	0.268	0.000	0.001	1.26
DOWJONES	0.530	0.000	0.000	0.475	0.407	0.478	1.06

Note: The different columns are: AS: Asymptotic test, SI: Sign test, WI: Wilcoxon's test, NB: Naive benchmark test, MGN: Morgan-Granger-Newbold's test, MR: Meese-Rogoff's test, n: number of times where the residuals coming from the method locating change-points in the t-ratios by minimizing the sum of squared residuals are smaller than the residuals coming from the spectral method, divided by the inverse number (when n>1, it means that the first method is better). The null hypothesis is the hypothesis of equal accuracy of different predictive methods. The loss function is quadratic. The test statistics follows asymptotically different distributions: $N(0,1)$ for the asymptotic test, the sign test, the Wilcoxon's test, the Meese-Rogoff's test, $F(T_0, T_0)$ for the Naive benchmark test and a t_{T_0-1} for the Morgan-Granger-Newbold's test (where T_0 is the number of predicted observations, i.e. $T_0 = 20$). The Meese-Rogoff test statistic is computed with the Diebold-Rudebusch covariance matrix estimator. The truncation lag is 2 for the asymptotic test and is given by the integer part of $T_0^{4/5}$ for the Meese-Rogoff's test.

Figure 1.- ACF and spectrum of the long-memory SETAR model



Simulated distribution of the bias ($\hat{c} - c$) over 100 replications

Chow test on t-ratios - d=0.40

Fig. 4a.
Case (i) - GPH estimator

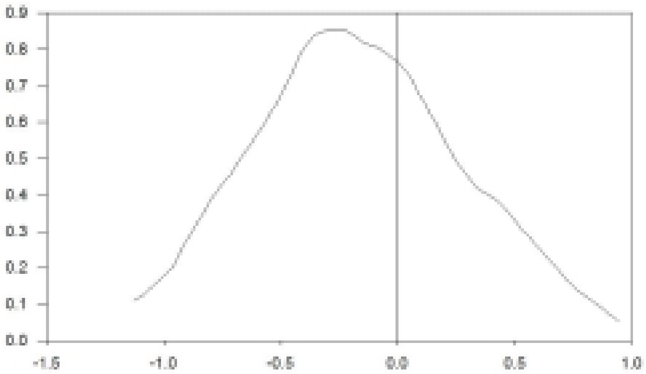


Fig. 4e.
Case (ii) - GPH estimator

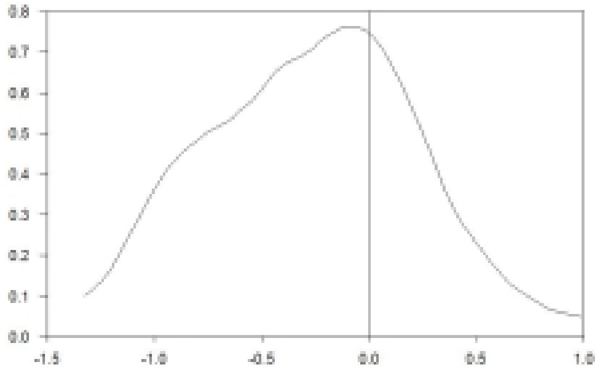


Fig. 4b.
Case (i) -Semiparametric estimator

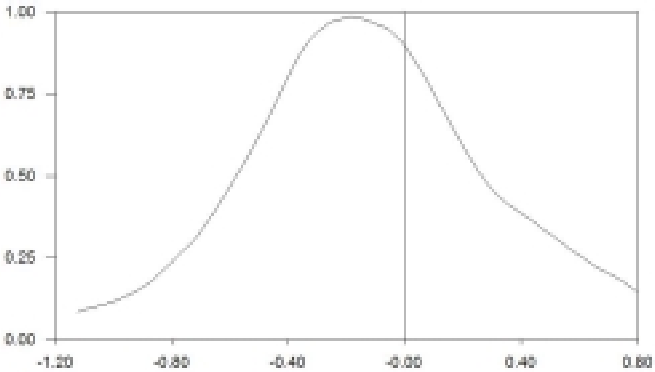


Fig. 4f.
Case (ii) - Semiparametric estimator

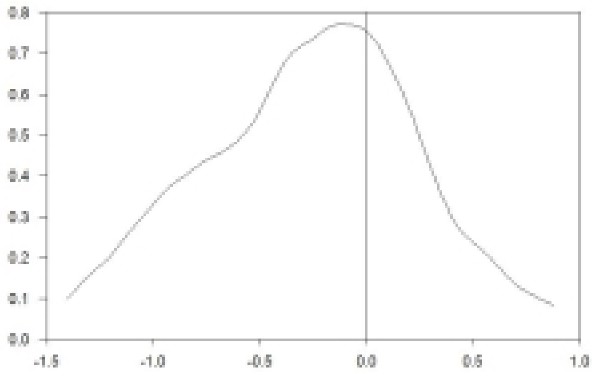


Fig. 4c.
Case (iii) - GPH estimator

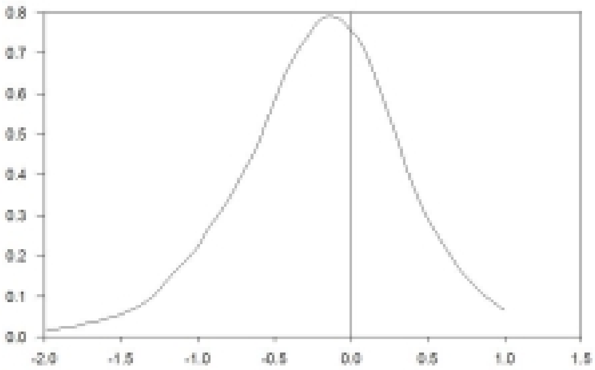
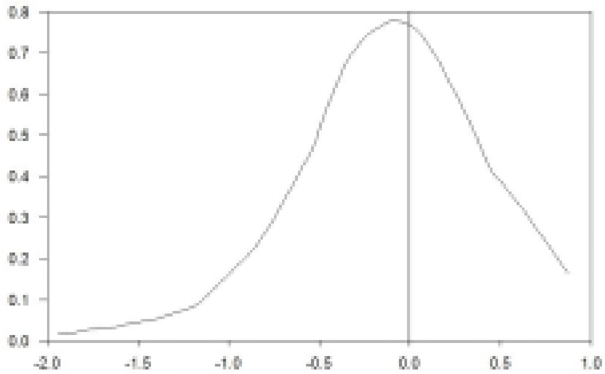


Fig. 4d.
Case (iii) - Semiparametric estimator



Simulated distribution of the bias ($\hat{c} - c$) over 100 replications

Minimization of the sum of squared residuals - d=0.49

Fig. 5a
Case (i) - GPH estimator

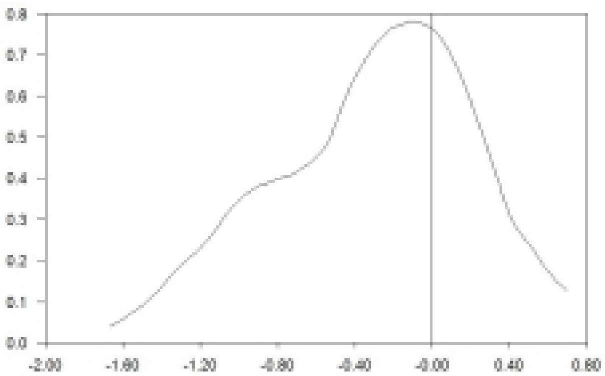


Fig. 5e
Case (ii) - GPH estimator

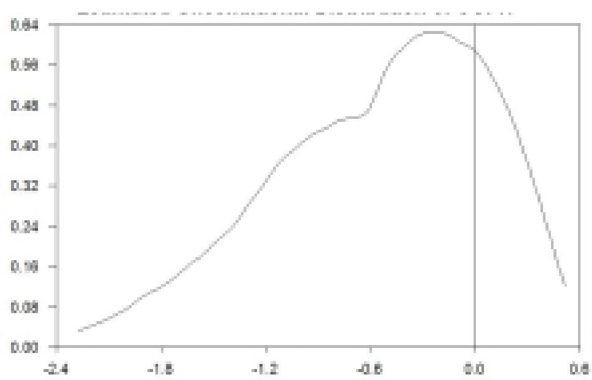


Fig. 5b
Case (i) - Semiparametric estimator

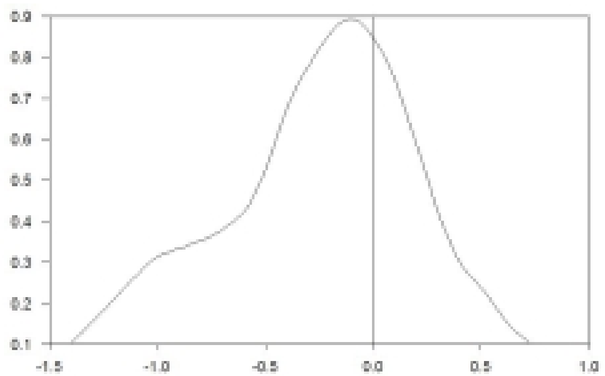


Fig. 5f
Case (ii) - Semiparametric estimator

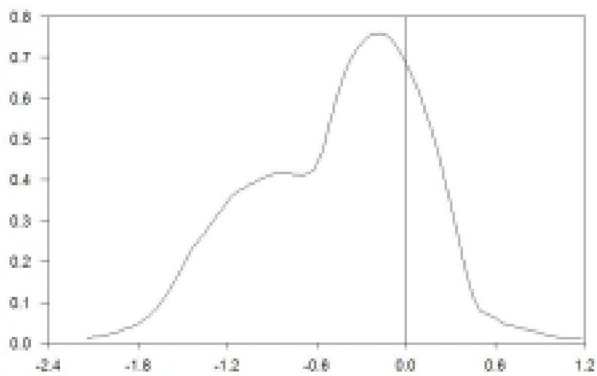


Fig. 5c
Case (iii) - GPH estimator

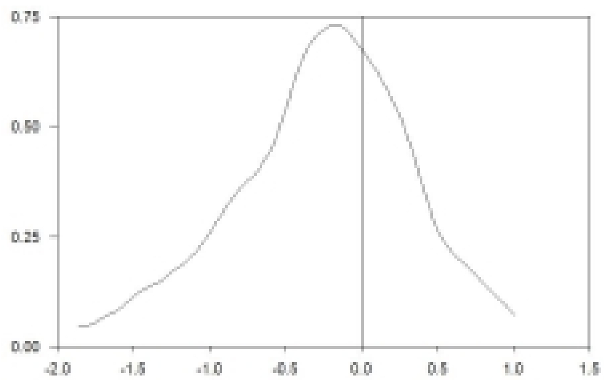
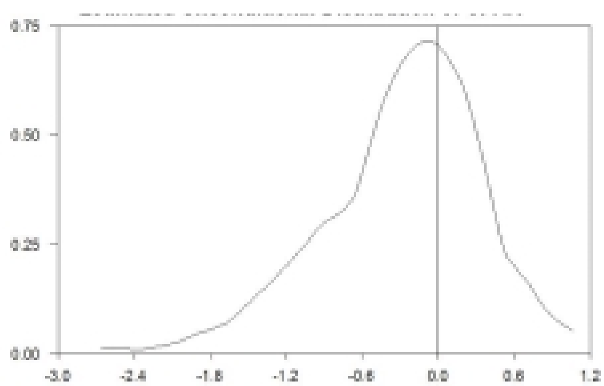


Fig. 5d
Case (iii) - Semiparametric estimator



Simulated distribution of the bias ($\hat{c}-c$) over 100 replications - Grid search method

Fig. 3a
d=0.49 - GPH estimator

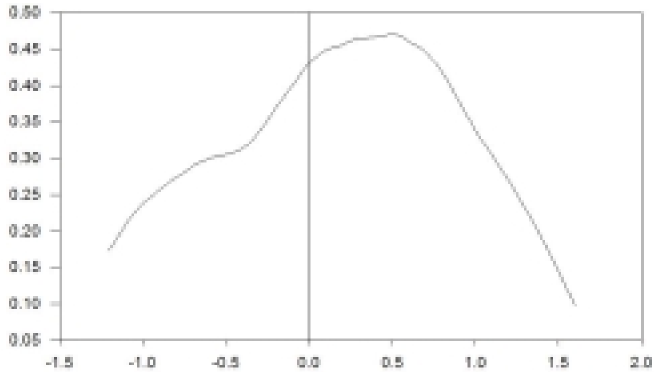


Fig. 3b
d=0.10 - GPH estimator

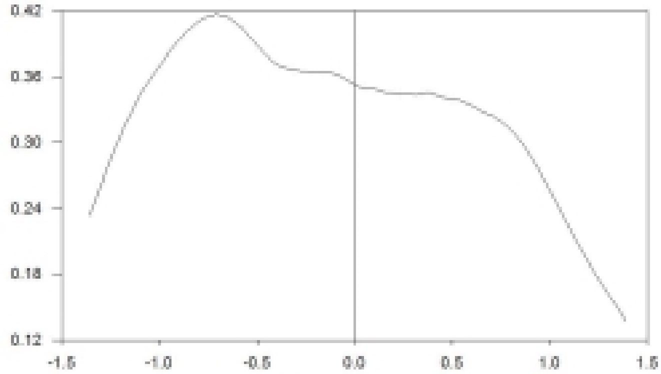


Fig. 3c
d=0.49 - Semiparametric estimator

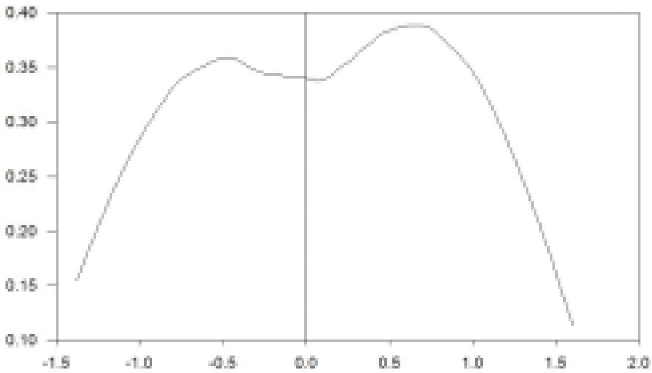


Fig. 3d
d=0.10 - Semiparametric estimator