



Department of Statistical Sciences  
*University of Padua*  
*Italy*

UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA  
DIPARTIMENTO  
DI SCIENZE  
STATISTICHE

## Spurious effects in switching regime processes

**Luisa Bisaglia**

Department of Statistical Sciences  
University of Padua  
Italy

**Margherita Gerolimetto**

Department of Statistical Sciences  
University of Padua  
Italy

**Abstract:** Long range dependence and regime switching are very intimately related effects. In this paper we consider the problem of spuriously detecting breaks point in hypothesis of long memory data generating processes. For this purpose we present some Monte Carlo evidence, together with a strategy to cope with the problem of distinguishing between long memory and regime switching.

**Keywords:** Long memory, Regime switching, break points

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>ARFIMA models</b>	<b>3</b>
<b>3</b>	<b>Structural breaks research</b>	<b>4</b>
3.1	Introduction . . . . .	4
3.2	The Bai-Perron test . . . . .	4
3.3	Models for structural breaks . . . . .	5
<b>4</b>	<b>Spurious effects</b>	<b>6</b>
4.1	Spurious long memory . . . . .	6
4.2	Spurious break points . . . . .	7
4.3	A strategy to deal with spurious effect . . . . .	8
<b>5</b>	<b>Finite sample experiment</b>	<b>9</b>
<b>6</b>	<b>Conclusions</b>	<b>11</b>

---

Department of Statistical Sciences  
Via Cesare Battisti, 241  
35121 Padova  
Italy

Corresponding author:  
Luisa Bisaglia  
tel: +39 049 827 4168  
luisa.bisaglia@unipd.it

tel: +39 049 8274168  
fax: +39 049 8274170  
<http://www.stat.unipd.it>

# Spurious effects in switching regime processes

**Luisa Bisaglia**

Department of Statistical Sciences  
University of Padua  
Italy

**Margherita Gerolimetto**

Department of Statistical Sciences  
University of Padua  
Italy

**Abstract:** Long range dependence and regime switching are very intimately related effects. In this paper we consider the problem of spuriously detecting breaks point in hypothesis of long memory data generating processes. For this purpose we present some Monte Carlo evidence, together with a strategy to cope with the problem of distinguishing between long memory and regime switching.

**Keywords:** Long memory, Regime switching, break points

## 1 Introduction

A rich literature on the econometrics of persistence has developed in recent years. At the beginning, the question of the nature of long term trend and consequent persistence in the empirical autocorrelation function had been explored by some researchers (among the others Nelson and Plosser, 1982) using ARIMA models or ARMA models around a deterministic trend. Others, for instance Harvey (1985), based their analysis on linear unobserved components models. A third choice employed the cointegrated specification of Engle and Granger (1987). All this interpretations were based on the assumption that first differences of the series follow a linear stationary process.

In the last fifteen years some alternative schemes have been proposed and considered to investigate the nature of persistence in economic time series. One approach has been studying mechanisms, such as structural breaks, that can somehow cause this form of long range dependence in the series. The idea is that the underlying DGP is a nonlinear process subjected to discrete shifts in regimes, that are episodes across which the dynamic behaviour of the series is markedly different. For example, the economy may either be in fast growth or in a slow growth phase, with the switch between the two governed by the outcome of Markov process (so called Markov Switching processes, Hamilton, 1989).

An alternative methodology for explaining persistence in time series has been to

look at their short or long memory properties. From an empirical perspective, long memory is related to a high degree of persistence of the observed data. The starting point of the literature on fractionally integrated processes has been the fact that many economic and financial time series show evidence of being neither  $I(1)$  nor  $I(0)$ , but more generally  $I(d)$ , where  $d$  is real. They show significant autocorrelation up to very long lags, often defined as hyperbolic rate. However, when first differenced those series appear as overdifferenced.

A significant strand of the debate has also considered the properties of tests for unit root, cointegration or long memory in the presence of structural breaks. It has been shown that persistence tests are severely compromised in terms of their size and power properties, in series which display breaks because these processes give the impression of persistence.

Recently, in several papers (among the others, Granger and Hyung, 2004; Diebold and Inoue, 2001), it has been pointed out that structural change can be easily confused with fractional integration since the structural breaks processes generate slowly decaying autocorrelations and other properties of  $I(d)$  processes. Often it happens that a time series exhibits long memory not because it is really  $I(d)$ , but because of the neglected occasional breaks in the series. As a consequence, level shifts in mean give rise to the observed long memory phenomenon and it is very difficult for almost all the procedures of estimation of  $d$  to actually recognize that in this specific case  $d = 0$  and, generally speaking, the estimate is spuriously set to  $d > 0$ .

From another point of view, the fractional integration of the DGP causes many breaks being detected spuriously by the usual estimation methods. Indeed, under assumptions of weakly dependent heterogenous error term, the consistency of a quasi maximum likelihood estimator or the least squares estimator of the breaks is well demonstrated. But, the estimators show quite different properties for the time series that is integrated by fraction or integer. One might find a break spuriously near the middle of the time series even though there is no break.

In this paper we propose a strategy to deal empirically with the spurious effects that can emerge when estimating both the parameter  $d$  and the break points in a series that exhibits some kind of persistence. Our idea is to estimate  $d$  and then repeat the estimation after filtering out the structural breaks that correspond to the points of estimated level shifts, since in hypothesis of no long memory, the process should be  $I(0)$  between breaks. If the estimation procedure still produces a value of  $\hat{d}$  that is bigger than zero, we can conclude that behind the persistence in the observed data, there should be a fractionally integrated DGP. Therefore the breaks points previously estimated are spurious. If, on the contrary the estimation of  $d$  is closer to zero, the long range dependence is caused by the occasional level shifts. Therefore what is spurious this time is  $d > 0$ .

To show the performance of this procedure we present some simulations results, whose starting points are some findings reported in Bisaglia and Gerolimetto (2005) about the performance of 5 different methods to estimate  $d$  in hypothesis of structural break processes. In that paper the authors found out that although the local Whittle estimate performs somehow better, generally speaking all methods encounter an evident difficulty in recognizing the structural break nature of the process. Especially for a relatively big number of expected breaks and a high value

of the jump size all methods produce values of  $\hat{d}$  far from zero.

In this paper the simulations will be realized in two phases. Firstly we will consider the case of some structural breaks processes where the procedure to estimate  $d$  wrongly produced  $\hat{d} > 0$  and again estimate  $d$  after filtering out the break points. This group of simulation puts in evidence the spurious long memory that appears because of the presence of some level shifts. Once they have been estimated and eliminated from the series, the latter should not look like a long fractionally integrated time series any more and  $\hat{d}$  should become closer to zero. Secondly we will consider some fractionally integrated DGP where  $d$  will be again estimated after filtering out the estimated breaks point. This second group of simulation shows how from  $I(d)$  DGP some break points are spuriously detected.

The plan of the paper is as follows. In Section 2, we recall some notions about ARFIMA process, in section 3 we briefly introduce some basics on the structural break research together with a description of the models we will use and most common techniques to estimate break points. Section 4 is devoted to discuss the spurious effects in structural breaks and fractionally integrated processes. In section 5 we describe our Monte Carlo experiment and Section 6 concludes.

## 2 ARFIMA models

The concept of long memory can be defined either in the time domain or in the frequency domain. In the time domain, a stationary discrete time series is said to be long memory if its autocorrelation function decays to zero like a power function. This means that the dependence between successive observations decays slowly since the observations are strongly correlated even when they are very far from each other. In the frequency domain, a stationary discrete time series is said to be long memory if its spectral density is unbounded at low frequencies.

The memory of a process can also be expressed in terms of the rate of growth of variances of partial sums,  $\text{var}(S_T) = O(T^{2d+1})$ , where  $(S_t) = \sum_{i=1}^t x_i$  and  $d$  is the long memory parameter. These definitions are not completely equivalent but there is a tight connection between them (Beran, 1994).

In this paper we consider one of the most popular long memory processes that is the Autoregressive Fractionally Integrated Moving Average process, ARFIMA( $p, d, q$ ) in the following, independently introduced by Granger and Joyeux (1980) and Hosking (1981). This process simply generalizes the usual ARIMA( $p, d, q$ ) process by assuming  $d$  to be fractional.

Let  $\epsilon_t$  be a white noise process having  $E[\epsilon_t^2] = \sigma^2$ . The process  $\{X_t, t \in \mathbf{Z}\}$  is said to be an ARFIMA( $p, d, q$ ) process with  $d \in (-1/2, 1/2)$ , if it is stationary and satisfies the difference equation

$$\Phi(B) \Delta(B) (X_t - \mu) = \Theta(B) \epsilon_t,$$

where  $\Phi(\cdot)$  and  $\Theta(\cdot)$  are polynomials in the backward shift operator  $B$  of degree  $p$  and  $q$ , respectively,  $\Delta(B) = (1 - B)^d = \sum_{j=0}^{\infty} \pi_j B^j$  with  $\pi_j = \Gamma(j - d) / [\Gamma(j + 1)\Gamma(-d)]$ , and  $\Gamma(\cdot)$  is the gamma function.

If  $p = q = 0$  the process  $\{X_t, t \in \mathbf{Z}\}$  is called Fractionally Integrated Noise and denoted by  $I(d)$ . When  $d \in (0, 1/2)$  the ARFIMA( $p, d, q$ ) process is stationary

and the autocorrelation function decays to zero hyperbolically at a rate  $O(k^{2d-1})$ , where  $k$  denotes the lag. In this case we say that the process has a long-memory behavior. When  $d \in (-1/2, 0)$  the ARFIMA( $p, d, q$ ) process is a stationary process with intermediate memory, when  $d \in (0, 1/2)$  the process is stationary, invertible and possesses long-range dependence. We will assume for convenience and without loss of generality that  $\sigma^2 = 1$  and  $\mu = 0$ .

### 3 Structural breaks research

In this section we will introduce some basics about the structural break research and we also present the Bai-Perron test for multiple breaks. Finally we will describe three well-known break processes we will use in the Monte Carlo experiment.

#### 3.1 Introduction

In 1960s and 1970s business cycles were commonly thought of as departures from a secular trend. The trend was believed to be deterministic, often linear, while departures were assumed to be stationary therefore transitory.

Nelson and Plosser (1982) were the first to observe that the secular component could possess a stochastic nature and many long US macroeconomic time series had been examined and were unable to reject the null hypothesis of a unit root against trend-stationary alternatives for most of the series considered. In two seminal works Rappoport and Reichlin (1989) and Perron (1989) argued that the majority of shocks to the key economic variables of any economy would be transitory and that only few events would have any permanent effect. They represents such shocks as breaks in the underlying deterministic trends. They also demonstrated that if structural breaks were present in the data generating process but not allowed for the specification of an econometric model, the analysis would be biased towards erroneous non rejection of the unit root hypothesis.

The subsequent development of the literature took place in two directions. One refers to a scheme where, firstly the structural break point is estimated and then the model is built (Perron (1989) for the case of one break point, Bai (1997), Bai and Perron (1998) for the case of multiple breaks). The second direction refers to a scheme where the break points are interpreted as endogenous (Hansen, 2000).

In the first approach, the multiple breaks option is obviously more general since it allows for the possibility of multiple changes during the lifetime of the sample. Indeed, even when the date of the break is unknown, it is restrictive to allow only for one structural break. Just as a model of a process with a change of parameters would be misspecified if the presence of the break would be ignored, allowing for one break, when in fact multiple breaks are present, could lead to false conclusions.

#### 3.2 The Bai-Perron test

In this framework it is inserted the Bai and Perron (1998) procedure (hereafter BP) to estimate and test for multiple breaks at unknown dates. They considered a type test for the null hypothesis of no change versus a prespecified number of changes

and also an alternative of an arbitrary number of changes (up to some maximum) as well as a procedure that allows one to test for the null hypothesis of, say,  $l$  changes against the alternative hypothesis of  $l + 1$  changes. The latter is particularly useful in that it allows a specific to general modeling strategy to consistently determine the appropriate number of changes in the data. The tests can be constructed allowing different serial correlation in the errors, different distributions for the data and the errors across segments or imposing a common structure.

The relevant asymptotic distributions depend on a trimming parameter  $\epsilon = h/T$ , where  $T$  is the sample size and  $h$  is the minimal permissible length of a segment.

The model considered is a multiple regression model with  $m$  breaks ( $m + 1$  regimes)

$$y_t = x_t' \beta + z_t' \delta_j + u_t, \quad t = T_{j-1} + 1, \dots, T_j, \quad (1)$$

for  $j = 1, \dots, m + 1$ . In this model,  $y_t$  is the observed dependent variable at time  $t$ ,  $x_t$  ( $p \times 1$ ) and  $z_t$  ( $q \times 1$ ) are vectors of independent variables and  $\beta$  and  $\delta_j$ ,  $j = 1, \dots, m + 1$  are the corresponding vectors of coefficients,  $u_t$  is the disturbance at time  $t$ . The indices  $(T_1, \dots, T_m)$ , or break points, are treated as unknown ( $T_0 = 0, T_{m+1} = T$ ). This is a partial structural change model since  $\beta$  is not subject to the shift, whereas if  $p = 0$ , we would obtain a pure structural change model where all the coefficients vary with the break point.

The purpose is to estimate consistently  $(\beta^0, \delta_1^0, \dots, \delta_m^0, T_1^0, \dots, T_m^0)$  then to test for the presence of structural change. The method of estimation is based on the least squares principle, where for each  $m$ -partition of the sample  $(T_1, \dots, T_m)$ , estimates of the parameters are obtained by minimizing the sum of squared residuals (SSR) from Eq. 1. The estimated break points  $(\hat{T}_1, \dots, \hat{T}_m)$  are given by the outcome of the algorithm

$$(\hat{T}_1, \dots, \hat{T}_m) = \operatorname{argmin}_{T_1, \dots, T_m} S_T(T_1, \dots, T_m) \quad (2)$$

where  $S_T(T_1, \dots, T_m)$  is the SSR from Eq. 1 for a given  $m$ -partition and the minimization is taken over all partitions  $(T_1, \dots, T_m)$ .

### 3.3 Models for structural breaks

Now we briefly recall the structural breaks model we will consider in the paper. In all the models we are describing there are only occasional breaks in mean, which means that the number of breaks that can occur in a specific period of time is somehow bounded.

More formally, we assume that the probability of breaks,  $p$ , converges to zero slowly as the sample size increases, i.e.  $p \rightarrow 0$  as  $T \rightarrow \infty$ , yet  $\lim_{T \rightarrow \infty} Tp$  is a non-zero finite constant.

The first model we consider and describe is the so called mean plus noise or occasional break model (Chen and Tiao, 1990; Engle and Smith, 1999)

$$y_t = m_t + \epsilon_t, \quad t = 1, \dots, T, \quad (3)$$

where  $\epsilon_t$  is a noise variable and occasional level shifts,  $m_t$ , are controlled by two variables  $q_t$  (date of breaks) and  $\eta_t$  (size of jump), as

$$m_t = m_{t-1} + q_t \eta_t \quad (4)$$

where  $\eta_t$  is i.i.d.  $(0, \sigma_\eta^2)$ . In the following sections the distribution of  $\eta_t$  has been taken to be normal although this distribution has no particular relevance. We assume that  $q_t$  follows an i.i.d. binomial distribution, that is,

$$q_t = \begin{cases} 0, & \text{with probability } 1 - p \\ 1, & \text{with probability } p \end{cases} \quad (5)$$

The binomial model we described is so characterized by sudden changes only. It might be also the case that structural changes occur gradually. In this case a Markov Switching Model (Hamilton, 1989) is more adapt. Suppose  $s_t$  is a latent random variable that can assume only the two discrete values 0 or 1. Each value of  $s_t$  represents a different state in the length of memory of shock.  $s_t$  is assumed to be governed by the following Markov probability law:  $p_{ij} = Pr(s_t = j / s_{t-1} = i)$ . Then, it is possible to use a switching model of  $q_t$  such that  $q_t = 0$  when  $s_t = 0$  and  $q_t = 1$  when  $s_t = 1$ . In this specification a regime with  $s_t = 1$  represents a period of structural change. Therefore, there is structural change everytime  $s_t = 1$ , both if  $s_{t-1} = 1$  and if  $s_{t-1} = 0$ .

Finally we present the so called Stochastic Permanent Break Model (STOP-BREAK model) formulated by Engle and Smith (1999) to bridge the gap between transience and permanence of the shocks. The STOPBREAK is a stochastic process in which the long run impact of each observation is time varying and stochastic.

The formulation is as follows:

$$y_t = m_t + \epsilon_t \quad (6)$$

$$m_t = m_{t-1} + q_{t-1} \epsilon_{t-1} \quad (7)$$

where  $q_t = q(|\epsilon_t|)$  is non decreasing in  $|\epsilon_t|$  and bounded by zero and one, so that bigger innovations have more permanent effects, and  $\epsilon_t$  are i.i.d  $N(0, \sigma_\epsilon^2)$ , moreover  $q_t = \frac{\epsilon_t^2}{(\gamma + \epsilon_t^2)}$  for  $\gamma > 0$ . Therefore in the STOPBREAK process permanent shocks can be indentified by their larger magnitude. In this approach the effects of shocks can fluctuate between transitory and permanent and typically such data exhibit periods of apparent stationarity punctuated by occasional mean shifts.

## 4 Spurious effects

In this section we consider the problem of some spurious effect that can emerge in hypothesis of structural breaks DGP and long memory DGP because of the similarity of the two classes of processes.

### 4.1 Spurious long memory

The notion of long memory appears in many different fields, including hydrology, internet traffic, economic and finance. Observing long memory characteristics in a



given data set, however, does not imply necessary an underlying long range dependent process since there are spurious cases of this characteristic such as structural breaks. Granger and Terasvirta (1999), Granger and Hyung (2004), Diebold and Inoue (2001) demonstrate this fact using analytic and simulation evidence. Unfortunately, detecting the nature of long range dependence and differentiating between true and spurious nature of this dependence is a rather difficult task.

It has been shown recently that inference on the long memory parameter and persistence tests can be severely compromised in series which display occasional breaks, since these processes give the impression of persistence and their autocorrelation decays very slowly. This means that neglecting structural breaks causes an over estimation of the long memory parameter, leading to believe in a long memory data generating process.

Some part of the literature has been devoted to analyse the effects on the methods to estimate the long memory parameter on hypothesis of structural breaks. Recently Bisaglia and Gerolimetto (2005) carried out a simulation experiment to show that all the major methods to estimate  $d$  are somehow biased in hypothesis of occasional break data generating process. This means that if an occasional break process (for instance a STOBREAK process) is generated and the long memory parameter,  $d$ , is estimated, it is likely that the estimation method does not recognize the process as short memory with breaks in mean, but it produces instead a value of  $\hat{d}$  bigger than zero.

## 4.2 Spurious break points

Recently, some implications of long memory dependence when testing for structural change in a time series have been explored. The problem seems natural one to study because long memory time series analysis is often applied to series which extend over a long period of time, and the longer is the time period, the greater the possibility of structural breaks.

Hidalgo and Robinson (1996) proposed a test for a change in parameter values at a given point in linear regression models with long memory errors. Although the structural change test they proposed is designed specifically for  $I(d)$  data, they put in evidence that their test may still have large size distortions in small samples leading to estimate the change point even when there is none.

In another paper Kuan and Hsu (1998) show that in hypothesis of fractionally integrated data the least squares estimator of the change point is consistent only when a mean change occurs and that the rate of convergence depends on the value of  $d$ . Moreover, when the data have strong dependence, a much larger sample would be needed to ensure that the least squares estimator is as precise as that for data with short memory. On the contrary when there is no mean change, the least squares estimator seems to be inconsistent in hypothesis of positive  $d$ , suggesting in that way a spurious change point in the middle of the sample.

Granger and Hyung (2004) also demonstrated that, in case of a DGP with long memory characteristics, the estimation techniques, as for instance quasi-maximum likelihood or least squares, usually used to estimate break points encounter some difficulties in distinguishing between structural change and long memory. In particular,

as the value of the long memory parameter  $d$  increases, it grows also the estimated number of breaks even when the DGP have no break points. So the long memory properties of the DGP might cause many breaks to be detected spuriously by standard estimation methods and unlike  $I(0)$  processes, caution should be exercised in estimating break points in presence of  $d > 0$ .

These results confirm that spurious change may arise from stationary data with long memory. Hence nonstationarity is not a necessary condition for a spurious change and also distinguishing between long memory time series and short memory time series with structural change is as difficult as distinguishing between  $I(1)$  time series and  $I(0)$  time series with one or more break points.

### 4.3 A strategy to deal with spurious effect

In this section we suggest a strategy to cope with the problem of the spurious effects described above. When an estimation method produces  $\hat{d}$  bigger than zero, there might be sometimes the suspicion that this result is caused by spurious instead of true long memory. If this is the case, our idea to concretely understand the nature of this long range dependence is to repeat the estimation procedure of  $d$  in the series, where break points have been estimated and then if eliminated from the series itself.

If after filtering out the break points,  $\hat{d}$  is still bigger than zero, this means that long memory should be a true feature of the series and the structural breaks had been spuriously detected. This is consistent with the theory described in section 4.2 and what happens is that the long memory properties compromised the performance of the BP test, which for this reason estimated spuriously some break points in the series.

If on the contrary, after deleting from the series the break points,  $\hat{d}$  is closer to zero we can conclude that the long memory feature had been spuriously estimated at the beginning and was caused by the presence of level shifts giving the series the impression of persistence.

In summary, the following is a simple scheme of the procedure applied to a series that exhibits some form of persistence:

- Estimate  $d$  in the original series  $x_t$ , obtaining  $\hat{d}$
- Identify the break dates in  $x_t$
- Obtain the break-free series  $x'_t = x_t - \hat{m}_t$  where  $\hat{m}_t$  is the sample mean of  $x_t$  of each regime
- Estimate  $d$  in  $x'_t$ , obtaining  $\hat{d}'$

If  $\hat{d} \cong \hat{d}'$ , we conclude that  $x_t$  is a true long memory time series, whereas if  $\hat{d} \neq \hat{d}'$  and  $\hat{d}'$  goes closer to zero, the conclusion is that  $x_t$  should be characterized by structural breaks and not by long memory.

## 5 Finite sample experiment

In this section we describe our Monte Carlo study. It has been realized to show the performance of our procedure to distinguish between long memory and structural break and it consists of two groups of simulations. Firstly we focus our attention on spurious long memory by generating three structural break models and then we orient our interest to spurious structural breaks by generating ARFIMA models. The functions we use are written in R language (Ihaka and Gentleman, 1996) and are available upon request by the authors.

In the first part of the experiment we simulate:

1. DGP1: mean plus noise model (3), with  $p = 0.01, 0.05, 0.1$  and  $\sigma_\eta^2 = 0.01, 0.05, 0.1$ ;
2. DGP2: Markov switching model, with  $(p, q) = (0.95, 0.95; 0.95, 0.99; 0.99, 0.95; 0.99, 0.99; 0.999, 0.999)$  and  $\sigma_\eta^2 = 0.1$ . In this case the initial state  $s_1$  is generated by a Bernoulli random variable with  $p = 0.5$ ;
3. DGP3: STOPBREAK model (6), with  $\gamma = (10^{-5}, 10^{-1}, 1, 10, 10^3)$ .

For each model we have considered  $\sigma_\epsilon^2 = 1$  and  $s = 1000$  independent realizations. Thus for a given estimation method we obtain  $s = 1000$  estimated values for  $d$ . Moreover, to evaluate the effects of different sample size, we have considered  $T = 500, 1000$ . All series are generated with 200 additional values in order to obtain random starting values.

Since the objective of this Monte Carlo experiment is to evaluate the performance of our strategy we decided to put ourselves in the most difficult circumstances to distinguish between long memory and structural breaks. Therefore, following some results obtained in Bisaglia and Gerolimetto (2005), we choose for this experiment high values both in terms of expected number of breaks and size of the jumps. To estimate  $d$  we employed the 5 methods that had been considered by Bisaglia and Gerolimetto (2005): the R/S method (Hurst, 1951), the Higuchi method (Higuchi, 1988), the aggregate variance method, the log-periodogram or GPH method (Geweke and Porter-Hudak, 1983) and Whittle's pseudo-likelihood method (Fox and Taqqu, 1986).

In the second part of the experiment we simulate ARFIMA(0,  $d$ , 0) processes, where  $d = 0.1, 0.2, 0.3, 0.4, 0.45$  and the noise is normally distributed with zero mean and unit variance. Also in this second part  $d$  has been estimated with the five methods listed above.

Following the steps described in section 4.3 the procedure can be implemented by estimating  $d$  (obtaining  $\hat{d}$ ), identifying the break dates with the BP method, obtaining the break-free series and finally re-estimating  $d$  in the new series (obtaining  $\hat{d}'$ ). In the first group of simulations we expect to find  $\hat{d}'$  much closer to zero than  $\hat{d}$  since the DGP possesses structural breaks and the long memory effect is spurious. In the second group of simulations, we expect to find  $\hat{d}'$  not very far from the previous estimate  $\hat{d}$  because the DGP is long memory and the structural breaks detected by BP are spurious.

In the following tables the results are reported. In tables 1 – 6 the results of the first group of simulations are presented, i.e. the estimates and standard deviation (in parenthesis) of  $d$  for DGP1, DGP2, DGP3 relatively to the break-free series. If we compare the results with the corresponding estimation in the original series (Bisaglia and Gerolimetto, 2005) it is clear that once the level shift is removed from the series, the estimation of  $d$  are dramatically different.

As expected, by eliminating the break components from the series, the cause of its persistent appearance is reduced and for this reason the estimate of  $d$  is now smaller.

If we look at the results relative to DGP1 (table 1 and 2), all methods perform well since once the level shift is removed from the series the estimation of  $d$ ,  $\hat{d}'$ , goes close to zero meaning that the true nature of the series is definitely not long memory, but stationary with structural breaks. The good performance of all method is particularly striking in case of big jumps (high value of  $\sigma$ ) and several expected break points (high value of  $p$ ).

If we observe results in table 3 and 4 (DGP2) we can again conclude that the elimination of the breaks makes possible to recognize that the process is not long memory. Indeed, the first four methods produce negative  $\hat{d}'$  evidencing some form of overdifferencing (we remind that  $d < 0$  means antipersistence in the process), whereas Whittle's method produces  $\hat{d}'$  much smaller than in the original series although slightly bigger than zero. For this DGP the results are particularly interesting since in the original series (table 4 and 5, Bisaglia and Gerolimetto, 2005) the performance of all methods (including Whittle) were particularly bad.

Some more comments have to be made for table 5 and 6 (DGP3) where the improvement in the estimation of  $d$  after removing the breaks is not so evident as in the two previous cases. The reason of this only small improvement is probably the nature of the STOPBREAK process. Indeed in that process  $q_t$  is  $\frac{\epsilon_t}{\gamma + \epsilon_t}$  therefore if  $\gamma \rightarrow 0$ ,  $q_t \rightarrow 1$  and the process is almost  $I(1)$ , whereas if  $\gamma \rightarrow \infty$ ,  $q_t \rightarrow 0$  and the process is  $I(0)$ . As a consequence, when  $\gamma$  does not tend to zero all methods work, since the process is  $I(0)$  and it is actually easier to recognize the non long memory features. It is different when  $\gamma$  tends to zero, since the increasing similarity to an  $I(1)$  process complicates distinguishing it from a non stationary process. For the smallest values of  $\gamma$  the process is so close to  $I(1)$  that by differencing it any negative moving average components appear (results available by the authors upon request). It seems that in this case the strategy of differencing may be more adequate than looking for structural breaks.

A possible explanation of the better performance in DGP1 than in DGP2 and DGP3 could be that the latter two processes are characterized by structural breaks that arise in a somehow gradual way. Therefore it is more difficult to identify the actual date of the break. As a consequence also the estimation of  $d$  is affected by this weakness.

From a general point of view we observe that the performance of the suggested method is slightly less good with the increase of the sample size. As pointed out by Bisaglia and Gerolimetto (2005) this is consistent with the nature of the structural break processes, since the longer is the process the more level shifts can take place, so it is even more difficult to distinguish it from a long memory one.

**Table 1:** Estimation results for  $d$ : DGP1, T=500

$\sigma_\eta^2$		$p = 0.01$	$p = 0.05$	$p = 0.1$
0.01	rs	0.063 (0.089)	0.067 (0.095)	0.074 (0.093)
	av	-0.002 (0.078)	0.011 (0.079)	0.021 (0.083)
	hi	0.126 (0.178)	0.094 (0.216)	0.065 (0.210)
	gph	0.038 (0.178)	0.043 (0.176)	0.048 (0.188)
	wh	0.014 (0.021)	0.015 (0.021)	0.017 (0.022)
0.05	rs	0.057 (0.099)	0.054 (0.093)	0.044 (0.084)
	av	0.001 (0.085)	0.012 (0.082)	0.011 (0.082)
	hi	0.084 (0.218)	-0.028 (0.157)	-0.064 (0.122)
	gph	0.029 (0.187)	0.036 (0.186)	0.006 (0.199)
	wh	0.013 (0.020)	0.019 (0.024)	0.022 (0.026)
0.1	rs	0.056 (0.098)	0.046 (0.089)	0.027 (0.082)
	av	-0.002 (0.087)	0.011 (0.088)	-0.004 (0.085)
	hi	0.053 (0.216)	-0.057 (0.133)	-0.103 (0.106)
	gph	0.030 (0.186)	0.014 (0.195)	-0.031 (0.191)
	wh	0.012 (0.018)	0.021 (0.026)	0.027 (0.029)

Another interesting general comment is on the performance of the estimation methods of  $d$  inside the strategy. The best and most recommended method is definitely the Whittle method which is somehow the most robust over all the considered DGP. This result is very important especially if we take into account that we are considering the most difficult circumstances. The R/S method exhibits a good performance for the third DGP. The worst methods seem to be Higuchi and GPH, as already pointed out in the literature.

As far as the  $I(d)$  processes, in tables 7 and 8 the results are reported, both before and after removing the break components from the series. In the first line of the tables there are the estimation results for the original series. The second line contains the results for the break-free series. It seems that with the Whittle method our strategy works well also with the increasing of the sample size. The estimation of  $d$  produced in the final step is approximately stable around the value obtained in the first step, meaning that the procedure recognizes the true long memory nature of the process. We have to observe that there is a slight tendency to underestimate  $d$  in the second step, but this is expected because implementing the procedure means substrating the sample means of each estimated regime from the series and so somehow compressing the series and reducing his appearance of persistence. However when we use Whittle method this is only a negligible effect, whereas for the other estimation methods this phenomenon affects much more the estimates that turn up to be often negative, typical synthom of overdifferencing. As far as the GPH method is concerned, the results are consistent with the findings of Granger and Hyung (2004).

## 6 Conclusions

In this article we present a strategy to distinguish between a time series exhibiting long-range dependence and one with short memory but suffering from structural shifts. In particular, we focus our attention on the problem of the spurious effects

**Table 2:** Estimation results for  $d$ : DGP1, T=1000

$\sigma_\eta^2$		$p = 0.01$	$p = 0.05$	$p = 0.1$
0.01	rs	0.074 (0.085)	0.075 (0.088)	0.091 (0.079)
	av	0.023 (0.072)	0.041 (0.077)	0.078 (0.075)
	hi	0.112 (0.186)	0.028 (0.177)	0.019 (0.154)
	gph	0.061 (0.141)	0.083 (0.149)	0.132 (0.149)
	wh	0.011 (0.014)	0.016 (0.018)	0.029 (0.024)
0.05	rs	0.067 (0.093)	0.054 (0.081)	0.049 (0.071)
	av	0.027 (0.083)	0.043 (0.077)	0.049 (0.075)
	hi	0.021 (0.174)	-0.062 (0.106)	-0.094 (0.086)
	gph	0.063 (0.158)	0.079 (0.156)	0.079 (0.158)
	wh	0.015 (0.018)	0.036 (0.025)	0.058 (0.030)
0.1	rs	0.067 (0.089)	0.044 (0.064)	0.055 (0.059)
	av	0.031 (0.078)	0.043 (0.073)	0.056 (0.074)
	hi	-0.004 (0.148)	-0.101 (0.084)	-0.113 (0.073)
	gph	0.069 (0.155)	0.068 (0.154)	0.069 (0.147)
	wh	0.017 (0.021)	0.057 (0.033)	0.097 (0.036)

**Table 3:** Estimation results for  $d$ : DGP2, T=500

$\sigma_\eta^2$		$p = 0.95$ $q = 0.95$	$p = 0.95$ $q = 0.99$	$p = 0.99$ $q = 0.95$	$p = 0.99$ $q = 0.99$	$p = 0.999$ $q = 0.999$
0.1	rs	-0.031 (0.068)	0.015 (0.084)	-0.064 (0.055)	-0.025 (0.081)	-0.027 (0.079)
	av	-0.082 (0.088)	-0.042 (0.084)	-0.121 (0.079)	-0.085 (0.087)	-0.084 (0.091)
	hi	-0.185 (0.090)	-0.084 (0.134)	-0.236 (0.067)	-0.154 (0.121)	-0.132 (0.180)
	gph	-0.233 (0.220)	-0.064 (0.196)	-0.366 (0.192)	-0.221 (0.213)	-0.221 (0.247)
	wh	0.053 (0.039)	0.021 (0.026)	0.098 (0.049)	0.051 (0.046)	0.066 (0.059)

**Table 4:** Estimation results for  $d$ : DGP2, T=1000

$\sigma_\eta^2$		$p = 0.95$ $q = 0.95$	$p = 0.95$ $q = 0.99$	$p = 0.99$ $q = 0.95$	$p = 0.99$ $q = 0.99$	$p = 0.999$ $q = 0.999$
0.1	rs	-0.069 (0.053)	-0.018 (0.076)	-0.091 (0.038)	-0.046 (0.063)	-0.038 (0.076)
	av	-0.139 (0.072)	-0.060 (0.079)	-0.177 (0.058)	-0.117 (0.074)	-0.120 (0.089)
	hi	-0.246 (0.064)	-0.142 (0.098)	-0.281 (0.049)	-0.200 (0.089)	-0.156 (0.167)
	gph	-0.336 (0.159)	-0.126 (0.177)	-0.451 (0.140)	-0.284 (0.174)	-0.266 (0.228)
	wh	0.099 (0.041)	0.031 (0.028)	0.164 (0.037)	0.102 (0.046)	0.099 (0.073)

**Table 5:** Estimation results for  $d$ : DGP3, T=500

$\gamma$	$10^{-5}$	$10^{-1}$	1	10	$10^3$
rs	-0.093 (0.043)	-0.092 (0.041)	-0.089 (0.046)	-0.059 (0.060)	0.037 (0.086)
av	-0.204 (0.058)	-0.207 (0.058)	-0.192 (0.060)	-0.116 (0.084)	-0.018 (0.074)
hi	-0.279 (0.042)	-0.282 (0.042)	-0.278 (0.042)	-0.230 (0.074)	0.069 (0.131)
gph	-0.605 (0.162)	-0.589 (0.152)	-0.589 (0.152)	-0.345 (0.192)	-0.004 (0.176)
wh	0.695 (0.051)	0.657 (0.051)	0.472 (0.056)	0.080 (0.046)	0.009 (0.017)

**Table 6:** Estimation results for  $d$ : DGP3, T=1000

$\gamma$	$10^{-5}$	$10^{-1}$	1	10	$10^3$
rs	-0.057 (0.046)	-0.055 (0.048)	-0.010 (0.045)	0.128 (0.060)	0.044 (0.081)
av	-0.163 (0.049)	-0.163 (0.053)	-0.104 (0.061)	0.111 (0.069)	-0.001 (0.068)
hi	-0.271 (0.036)	-0.267 (0.039)	-0.229 (0.046)	-0.087 (0.068)	0.137 (0.155)
gph	-0.496 (0.107)	-0.505 (0.133)	-0.343 (0.128)	0.158 (0.141)	0.023 (0.146)
wh	0.825 (0.033)	0.793 (0.033)	0.656 (0.036)	0.324 (0.045)	0.009 (0.014)

**Table 7:** Estimation results for  $d$ : DGP ARFIMA(0,d,0), T=500

$d$	0.1	0.2	0.3	0.4	0.45
rs	0.117 (0.089) 0.075 (0.087)	0.201 (0.103) 0.048 (0.077)	0.272 (0.111) 0.000 (0.069)	0.360 (0.120) -0.033 (0.056)	0.393 (0.129) -0.041 (0.054)
av	0.068 (0.074) 0.024 (0.076)	0.153 (0.075) 0.011 (0.078)	0.216 (0.074) -0.037 (0.073)	0.289 (0.073) -0.080 (0.069)	0.317 (0.069) -0.095 (0.074)
hi	0.074 (0.107) 0.017 (0.122)	0.184 (0.109) -0.029 (0.147)	0.262 (0.117) -0.138 (0.089)	0.369 (0.104) -0.194 (0.072)	0.417 (0.098) -0.208 (0.071)
gph	0.092 (0.170) 0.020 (0.178)	0.216 (0.169) -0.029 (0.174)	0.292 (0.173) -0.179 (0.183)	0.404 (0.179) -0.305 (0.182)	0.463 (0.182) -0.335 (0.178)
wh	0.093 (0.036) 0.079 (0.037)	0.194 (0.033) 0.148 (0.041)	0.291 (0.036) 0.215 (0.044)	0.396 (0.039) 0.291 (0.046)	0.446 (0.036) 0.332 (0.047)

**Table 8:** Estimation results for  $d$ : DGP ARFIMA(0,d,0), T=1000

$d$	0.1	0.2	0.3	0.4	0.45
rs	0.101 (0.089) 0.062 (0.077)	0.199 (0.099) 0.021 (0.067)	0.286 (0.109) -0.017 (0.054)	0.354 (0.118) -0.034 (0.048)	0.399 (0.114) -0.032 (0.042)
av	0.063 (0.067) 0.027 (0.067)	0.155 (0.071) -0.007 (0.066)	0.235 (0.071) -0.056 (0.067)	0.292 (0.068) -0.089 (0.061)	0.325 (0.062) -0.099 (0.060)
hi	0.080 (0.111) 0.024 (0.118)	0.181 (0.110) -0.085 (0.122)	0.275 (0.106) -0.169 (0.079)	0.359 (0.109) -0.216 (0.062)	0.401 (0.105) -0.220 (0.060)
gph	0.086 (0.133) 0.039 (0.134)	0.198 (0.131) -0.056 (0.153)	0.307 (0.139) -0.163 (0.145)	0.406 (0.149) -0.246 (0.141)	0.454 (0.136) -0.269 (0.053)
wh	0.094 (0.025) 0.087 (0.024)	0.196 (0.025) 0.165 (0.027)	0.296 (0.027) 0.242 (0.031)	0.399 (0.027) 0.333 (0.032)	0.449 (0.026) 0.379 (0.032)

that can arise in hypothesis of processes with structural breaks and  $I(d)$  processes. In the first case the presence of level shifts make the process look like a long memory one, so when  $d$  is estimated the obtained value  $\hat{d}$  is generally far from zero although the process is short memory.

In the second case the presence of long memory compromises the consistency of the techniques to estimate the break points. This means that if one estimates the break points on a long memory process, one or more break points would probably be estimated although the process is  $I(d)$  and not an occasional break one.

Both these spurious effects create many difficulties to distinguish between long memory and structural break processes as already pointed out in some papers in literature (for example, Granger and Hyung, 2004).

The strategy we propose consists of estimating the long memory parameter  $d$  on a series that gives the impression of persistence, then detecting the structural break dates and finally re-estimating  $d$  on the series after filtering out the estimated break points. If the estimated  $d$  is still unchanged after removing the break components, this means that the series should be generated from a true long memory process, whereas if the estimated  $d$  changes dramatically becoming closer to zero, this means that the true process should be an occasional break process.

To show the performance of our strategy we realize a Monte Carlo experiment in two parts in order to show what happens both in case of long memory and occasional breaks. The latter has been considered in very difficult cases, i.e. big jumps and several expected break points. By observing the tables of our results, it seems that

generally speaking both in case of structural break and long memory the strategy manages to approximately recognize the true nature of the process, especially if the estimation of  $d$  is done with the Whittle method.

These results seem to be rather encouraging and to our knowledge any other procedures have not been proposed to distinguish between the occasional breaks and long memory features. However, we are aware that a formal test is needed and it is under way by the authors.

## References

- Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric Theory*, **13**, 315–352.
- Bai, J. (1998). A note on spurious break *Econometric Theory*, **14**, 663–669.
- Bai, J. and Perron, P. (1998) Estimating and testing linear models with multiple structural changes. *Econometrica*, **66**, 47–78.
- Beran, J. (1994). *Statistics for Long-Memory Processes*. Chapman and Hall, London.
- Bisaglia, L. and Gerolimetto M. (2005) Switching regime and ARFIMA processes *Working Paper N.3*. Dep. of Statistical Sciences, University of Padova.
- Chen, C. and Tiao, G.C. (1990). Random level-shift time series models, ARIMA approximations, and level-shift detection *Journal of Business and Economics Statistics*, **8**, 83–97.
- Diebold, F.X. and Inoue, A. (2001). Long memory and regime switching *Journal of Econometrics*, **105**, 131–159.
- Engle, R.F. and Granger, C.W.J. (1987). Co-integration and error correction: representation, estimation and testing *Econometrica*, **55**, 251–276.
- Engle, R.F. and Smith, A.D. (1999). Stochastic permanent breaks *The Review of Economics and Statistics*, **81**, 553–574.
- Fox, R. and Taqqu, M.S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, **14**, 517–532.
- Geweke J. and Porter-Hudack S. (1983) The Estimation and Application of Long-Memory Time Series Models *J.T.S.A.*, **4**, (1983) 221-237.
- Gourieux, C. and Jasiak, J. (2001). Memory and infrequent breaks. *Economics Letters*, **70**, 29–41.
- Granger, C.W.J. and Joyeux, R. (1980). An introduction to long-range time series models and fractional differencing. *Journal of Time Series Analysis*, **1**, 15–30.



- Granger, C.W.J. and Hyung, N. (2004). Occasional structural breaks and long memory with an application to the S&P500 absolute stock returns. *Journal of Empirical Finance*, **11**, 399–421.
- Granger, C.W.J. and Terasvirta, T. (1999). A simple nonlinear time series model with misleading linear properties. *Economics Letters*, **62**, 161–165.
- Hamilton, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, **57**, 357–384.
- Hansen, B.E. (2000) Testing for structural change in conditional models. *Journal of Econometrics*, **97**, 93–115.
- Harvey, A.C. (1985) Trends and cycles in macroeconomic time series. *Journal of Business and Economic Statistics*, **3**, 216–227.
- Hidalgo, J. and Robinson P.M. (1996) Testing for structural change in a long-memory environment. *Journal of Econometrics*, **70**, 159–174.
- Higuchi, T. (1988). Approach to an irregular time series on the basis of the fractal theory. *Physica D*, **31**, 277–283.
- Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*, **68**, 165–176.
- Hurst, H.E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, **116**, 770–799.
- Ihaka, R. and Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, **5**, 299–314.
- Kuan, C.M. and Hsu, C.C. (1998) Change point estimation of fractionally Integrated Processes. *Journal of Time Series Analysis*, **19**, 693–708.
- Mikosh, T. and Starica, C. (1999). Change of structure in financial time series, long range dependence and the GARCH model. *Working Paper*. The Wharton School, Chalmers University of Technology
- Nelson, C.R. and Plosser, C.I. (1982) Trends and random walks in macroeconomic time series. *Journal of Monetary Economics*, **10**, 139–162.
- Perron, P. (1989) The great crash, the oil shock and the unit root hypothesis. *Econometrica*, **57**, 1361–1401.
- Rappoport, P., Reichlin, L., (1989) Segmented trends and nonstationary time series. *The Econometric Journal*, **9** (Conference volume), 163–168.



**Working Paper Series**  
**Department of Statistical Sciences, University of Padua**

You may order paper copies of the working papers by emailing [wp@stat.unipd.it](mailto:wp@stat.unipd.it)

Most of the working papers can also be found at the following url: <http://wp.stat.unipd.it>

